Reduction of Information Collection Cost for Inferring Brain Model Relations From Profile Information Using Machine Learning

Ryoichi Shinkuma, *Senior Member, IEEE*, Satoshi Nishida, Naoya Maeda, Masataka Kado, and Shinji Nishimoto

***Abstract*—A content recommendation system based on human brain activity has become a reality. However, the cost of collecting the information from people is problematic. This article proposes a scheme that resolves the tradeoff between the inference performance from a profile model to a brain model and the cost of collecting profile information. In the proposed scheme, a machine learning model infers the brain model from the profile model and a feature selection method is applied to reduce the cost, i.e., the number of questionnaire items, of collecting profile information. Since only the top questionnaire items with the highest importance scores are used, we can maintain the inference performance as high as possible while limiting the number of questionnaire items. We demonstrate the effectiveness of the proposed scheme with a performance evaluation using an experimentally obtained brain model and a profile model created from real profile information. The results over different experimental parameters, video lengths, and feature selection methods demonstrate that the proposed scheme successfully identifies the top questionnaire items that contribute most significantly to the inference of brain models.**

***Index Terms*—Brain model, feature selection, machine learning (ML), profile information.**

I. Introduction

THE BRAIN of a person might know more about that person than the actual person does [1], [2]. On the basis of this hypothesis, a content recommendation system based on brain activity has been proposed [3]. In this system, a brain scanner, such as a functional magnetic resonance imaging (fMRI) scanner is used to acquire brain signals from a subject who is perceiving stimuli of a content and then, by increasing the number and variety of subjects and content, it is eventually possible to create a general model of brain activity that reflects a wide variety of people's perceptions for a wide variety of content. Several research efforts have contributed to the creation of a general model of brain activity applicable for content recommendation [4]–[8]. Once we obtain the "ultimate" brain model, we should be able to perfectly estimate the brain activity of a person who is perceiving the stimuli of a specific content. This can potentially provide powerful solutions in the commercial field. For example, since advertisements are typically expensive to produce, in terms of both monetary and system costs [9], determining which advertisements are more attractive to different consumers in advance by using a brain model would significantly reduce costs.

As a key step toward creating the ultimate brain model, we have previously developed a relational network of people (RNP) based on brain imaging [3]. This model represents the similarity of the brain activities of people who are watching a wide variety of video content. We found that two people close to each other in the RNP model would show similar brain activity when they watch the same video content (see the brief explanation in Section IV-A). Therefore, we can expect that this model will work well for clustering people who have similar video preferences. A drawback is the cost related to brain scanning, including: 1) the deployment and maintenance costs of fMRI; 2) the time and labor costs of MRI specialists; and 3) the recruitment and experimental costs of subjects. Therefore, we targeted the cost reduction of brain scanning as a new research issue and considered that inferring the brain model from the profile information could be a solution. However, collecting profile information can be problematic, as people are generally hesitant about the collection of their private information through things like questionnaires.

In this article, we propose a scheme that resolves the tradeoff between the inference performance from a profile model, which is created from profile information, to a brain model and the cost of collecting profile information. Our scheme utilizes a machine learning (ML) method to create a model that enables the brain model to be inferred from the profile model. To reduce the cost of collecting profile information, the proposed scheme takes advantage of the feature selection capability of ML. Specifically,

Manuscript received December 12, 2020; accepted April 3, 2021. Date of publication July 15, 2021; date of current version June 16, 2022. This work was supported in part by JSPS KAKENHI under Grant 18K18141, and in part by JST PRESTO under Grant JPMJPR20C6. This article was recommended by Associate Editor J. Lu. *(Corresponding author: Ryoichi Shinkuma.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by The ethics and safety committee of the National Institute of Information and Communications Technology. Performed in line with The Helsinki Declaration and the ethics guideline given by the Japanese government.

Ryoichi Shinkuma was with the Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan. He is now with the Faculty of Engineering, Shibaura Institute of Technology, Tokyo 135-8548, Japan (e-mail: shinkuma@shibaura-it.ac.jp).

Satoshi Nishida and Shinji Nishimoto are with the Center for Information and Neural Networks, National Institute of Information and Communications Technology, Suita 565-0871, Japan.

Naoya Maeda and Masataka Kado are with the Social Innovation Division, Social Infrastructure Solution Sector, NTT DATA Corporation, Tokyo 135-8671, Japan.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSMC.2021.3074069.

Digital Object Identifier 10.1109/TSMC.2021.3074069

when we use a questionnaire to collect profile information, feature selection is utilized to estimate the extent to which each questionnaire item contributes to the inference performance and to quantitatively determine the score for each item in accordance with its importance. By using only questionnaire items with the highest importance scores, the proposed scheme maintains the inference performance as high as possible while reducing the number of questionnaire items, which results in the suppression of the cost of profile information collection. This article validates the proposed scheme through an evaluation using the brain model obtained through an experiment in our previous work [3] (see also the brief explanation in Section IV-A) and profile information obtained for the creation of the profile model.

In Section II of this article, we present related work. Section III describes the system model and the problem formulation in this work, followed by the methodology of the proposed scheme. Section IV presents the evaluation method and Section V reports the results. We conclude in Section VI with a brief summary.

II. Related Work

A. Cross Relation Inferring

Recently, inferring the target relations from another relation as a source has been attracting research interest. This section presents prior works on this subject. However, as explained below, the objective of these prior works and the source and target relations they examined are completely different from the ones in this work.

A couple of pioneering works appeared around 2012 to 2015. Tang *et al.* [10] presented a framework that infers the labels of links in a target network from multiple source networks. They gave application examples showing that the framework can infer trust/distrust relationships from a reviewer network and social relationships, such as families, friendships, or colleagues from mobile communication networks. They demonstrated that the framework successfully inferred manager-subordinate relationships in an enterprise e-mail network from a coauthor network with labeled advisor-advisee relationships. Fang *et al.* [11] studied a transfer learning approach to infer a target network from a source network. They applied their approach to four real network datasetstwo citation networks, a hyperlink network, and a network of co-located terrorist attacksand showed that inference between all pairs of different networks could be achieved. Wang *et al.* [12] investigated the extraction of potential links between activity spaces in a city. Other studies have shown that social network-based resource allocation in wireless and wired networks is effective because a communication link is established between socially connected users, such as families, friends, or colleagues [13]–[17].

Due to rapid advancements with deep learning, more recent studies tend to use the graph neural network as an ML model for inferring a target network from a source network. Fan *et al.* [18] presented GraphRec, a graph neural network framework for social recommendations. An example application of GraphRec is inferring a user-item graph where the weights of edges denote the opinions (or the rating score) of users on the items from a social graph between users. Dai *et al.* [19] and Shen *et al.* [20] worked on the cross-network node classification problem where the labels of unlabeled nodes in a target network are inferred from the labels of nodes in a source network. They aimed to transfer knowledge from a partially labeled source-attributed network to assist the classification task in a completely unlabeled or partially labeled target-attributed network. Using three real-world datasets of paper citation networks produced from different original sources, in which nodes (papers) are labeled with five categories, they showed that the inferred labeling of their framework had good accuracy [19]. They further examined it using datasets other than citation networks, e.g., the social networks of bloggers [20].

B. Functional Brain Imaging and Individual Profiles

Most of the previous fMRI studies have sought to uncover generic aspects of brain functions by averaging brain data across individuals to improve the signal-to-noise ratio, which results in little attention to individual differences in brain information. However, there have been recent attempts to apply fMRI research to individual differences in brain information and the correlation of individual profiles [21].

The majority of these attempts have focused on the individual variability of functional connectivity, which is characterized by the correlation of brain activity between different brain regions. Van Den Heuvel *et al.* [22] demonstrated that the network efficacy of individual brains evaluated by functional connectivity is correlated with the individual difference of intellectual performance. Finn *et al.* [23] also found a relationship between functional connectivity and intellectual performance at the individual level. In addition, they showed that the individual patterns of functional connectivity can be used as a fingerprint to distinguish one brain from another. Other studies have reported correlations of functional connectivity with different types of individual profiles, such as five-factor personality profiles (Big Five) [24] and creative thinking ability [25]. These studies evaluated functional connectivity using brain activity during the resting state, in which individuals do nothing in an MRI scanner with no explicit task. There is a concern that such resting-state functional connectivity may fail to capture the full range of individual differences in brain information [26], [27]. For this reason, attempts have been made to evaluate the individual difference of functional connectivity by using task-induced brain activity [28]. In one of these attempts, Greene *et al.* [29] reported that task-based functional connectivity is more strongly correlated with intellectual performance at the individual level than resting-state functional connectivity.

Another line of study has quantified individual differences of brain information using brain activity evoked by natural stimuli. Parkinson *et al.* [30] evaluated the similarity of brain information between individuals using brain activity evoked by natural movies and showed that the brain-based similarity between arbitrary pairs of individuals can be used to predict the social distance between the same pairs. Charest *et al.* [31] evaluated brain representations of objects using the representational similarity of brain activity evoked by natural images of the objects and showed that individuals have unique brain representations for objects that are personally meaningful

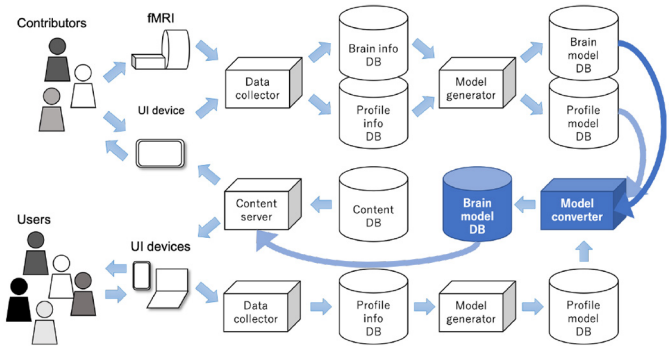


Fig. 1. System overview.

to themselves. Nishida and Nishimoto [4] developed an fMRI-based decoding method that recovers perceptual content induced by natural movies in the form of words and reported that the individual variability of the decoded content correlates with that of movie-scene descriptions. Shinkuma *et al.* [3] visualized brain-based relational networks of individuals using brain activity evoked by natural movies and reported that the relational networks are consistently obtained across datasets.

The present study differs from the related studies above in the following three aspects. First, while each of the above studies has focused on the fMRI correlations of one or two individual profiles only, the present study comprehensively seeks fMRI correlations of a wide range of profiles (209 items). Second, more importantly, while the related studies have basically aimed to predict individual profiles from brain activity, the present study ultimately aims to recover brain information from individual profiles for the sake of reducing fMRI measuring costs for brain-based content evaluation. Third, the aim of this study is not to develop new ML and feature selection methods but rather to propose a new scheme to reduce the information collection cost for inferring brain models from profile information. To this end, we applied conventional ML and feature selection methods to brain data and profile information in a unique way that has not been adopted by any existing studies. Thus, this study has novelty and innovation in terms of how to combine ML and feature selection methods with brain data and profile information.

III. Proposed Scheme

A. System Model

Fig. 1 shows the system model examined in this article. It consists of the fMRI, user interface (UI) devices, such as tablets, a content server, data collectors, model generators, a model converter, a content database (DB), profile information DBs, brain information DBs, profile model DBs, and brain model DBs.

Contributors are special users who allow the service provider to acquire their brain information by using the data collector with fMRI in advance and at regular intervals thereafter. The UI devices for contributors display content and questionnaires to acquire their brain and profile information, respectively. In this work, we examine short video content with audio (such as a TV advertisement), though this can be replaced with long video, music, or other content as future work. The acquired profile and brain information is stored in the profile and brain information DBs. The model generator generates the brain and profile models of each contributor and stores them in the profile and brain model DBs.

In contrast to the special users, normal users simply use UI devices for normal activities, such as browsing websites, purchasing products on e-commerce sites, watching videos, listening to music, and so on. The data collector acquires the profile information from these users via UI devices and stores it in the profile information DB. The model generator generates the profile model and stores it in the profile model DB. The model converter, which is a key component in the proposed scheme, learns the correlation between the brain models and the profile models of contributors and creates an ML model that enables the brain models for users to be inferred from their profile models. The inferred brain models are stored in the brain model DB and are used by the content server when a short video content likely to be "attractive" for a user is selected and presented to that user.

B. Problem Statement

This section presents the problem statement considered in this work. The system described in the previous section aims to maximize the quality of the content selection for users while keeping the total cost for the data collection below a certain threshold. The problem formulation of the system is given as

$$\max_{V,B,P} \quad Q_U(V, B, P), \tag{1}$$

$$\sum_{v\in V}(C_v(B) + C_v(P)) + \sum_{u\in U} C_u(P) \leq \hat{C} \tag{2}$$

where V, B, P, U, and $Q_U()$ indicate a set of contributors, a data collection method of brain information, a data collection method of profile information, a set of users, and the quality of the content selection for U, respectively. $C_x(Y)$ and $\hat{C}$ indicate the cost of a data collection method Y for person x and the total cost acceptable for the system. V, B, and P are decision variables in the problem formulation, while the others are given. Equation (1) is the objective of the system, and (2) is the constraint. We here describe the reason the constraint needs to be given as (2). As we mentioned in Section I, the costs for acquiring brain information and for collecting profile information are problematic; the former corresponds to $C_v(B)$, while the latter corresponds to $C_v(P)$ and $C_u(P)$. Optimizing the objective given as (1) without considering (2) means that the system does not care about those costs even if they are huge, which is an unrealistic assumption. Therefore, as presented in (2), the total cost given on the left side has to be limited by the acceptable cost of the system, $\hat{C}$.

C. Approach

Let us consider possible approaches for the problem formulation given in (1) and (2). One approach is to effectively recruit V so as to increase $Q_U(V, B, P)$ while suppressing the increases of $C_v(B)$ and $C_v(P)$ by keeping the size of V small. Another approach is to create a better brain model for B that increases $Q_U(V, B, P)$, as we explored in our previous work [3], while suppressing the increase of $C_v(B)$. Another

approach is to compose a P for questionnaire items that effectively increase $Q_U(V, B, P)$ while suppressing the increases of $C_v(P)$ and $C_u(P)$. In this article, we focus on the last approach. To clarify this research focus, in the following, we consider only P as a decision variable, while V and B are given parameters. More specifically, we assume that B is determined by the method presented in our previous work [3] (see also the brief explanation in Section IV-A) and that V is also fixed to the contributors recruited for our previous work. On the basis of these assumptions, we should be able to increase $Q_U(V, B, P)$ by improving the inference performance from the profile model to the brain model. Hence, we define a data collection method for the profile information, P, as a list of questionnaire items. We need to determine which questionnaire items contribute to improving the inference performance as well as to what extent each item contributes. To suppress the increases of $C_v(P)$ and $C_u(P)$, we need to limit the questionnaire items we use to only those that significantly contribute to the improvement.

D. Methodology of Proposed Scheme

This section presents the methodology of the proposed scheme, which is shown in Fig. 2. As we defined in the previous section, P is a list of questionnaire items. We also define A_v as a list of answers for the P of contributor v. Let D denote a matrix that represents differences of the profile information between all pairs of contributors. An element d_{12} in D is the difference of the profile information between contributors 1 and 2, A_1 and A_2. Suppose we have already obtained a matrix that represents the differences of the brain information between all pairs of contributors, Δ, in which an element δ_{12} is a real value and indicates the difference of the brain information between contributors 1 and 2. The Δ can be obtained by using the method presented in our previous work [3] (explained in Section IV-A). The D and Δ defined above are the profile and brain models mentioned in Section III-A, respectively. The model converter in the system learns the correlation between D and Δ and creates an ML model that infers the brain model for users, Δ', from the profile model for users, D'. In other words, the ML model created by the model converter has the matrices of the profile and brain models as its input and output, respectively.

In the next step right after the construction of the ML model (Fig. 2), the proposed scheme utilizes the feature selection capability of ML. Feature selection enables us to evaluate the contribution of each input variable to the inference performance of ML [32], [33]. In other words, it enables us to obtain the importance score of each input variable. In the proposed scheme, we use feature selection to estimate the extent to which each questionnaire item in P contributes to the inference improvement from the profile model, D, to the brain model, Δ, and an importance score is then given to each questionnaire item in accordance with the contribution. Then, we can select questionnaire items from the top in descending order of the importance scores in P. At the same time, by removing low-importance questionnaire items, we can keep the number of items smaller than a threshold, which results

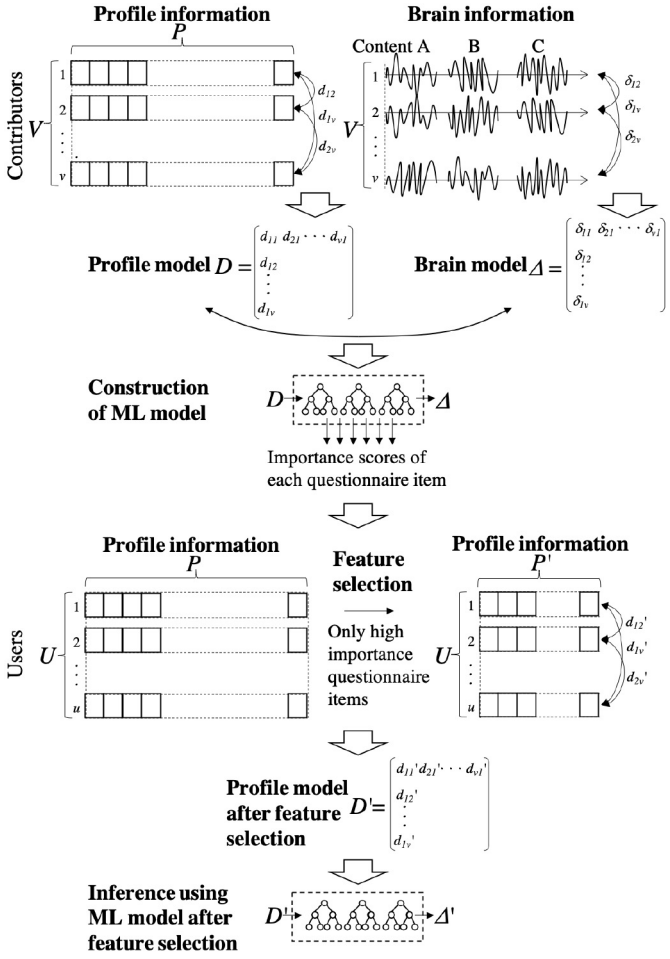


Fig. 2. Methodology of the proposed scheme.

in suppression of the costs [$C_v(P)$ and $C_u(P)$ in (2)]. Thus, the feature selection does not change the structure of the ML model of the model converter at all, yet it reduces the number of items used for calculating the correlation between each pair of users in the matrix of the profile model, D.

IV. Evaluation Method

This section presents the method we used to evaluate the proposed scheme. It consists of the following steps.

1) Recruitment of contributors.
2) Collection of brain information.
3) Creation of brain model.
4) Collection of profile information.
5) Creation of profile model.
6) Estimation of brain model from profile model using ML.
7) Extraction of importance score of questionnaire items using feature selection.

We explain each of these steps in detail below.

As stated above, we assumed the same contributors as the ones in our previous work [3]. Two contributor groups A and B consist of 26 and 39 contributors, respectively. They were recruited from among members of the public in a local area close to an experimental facility housing an MRI scanner in Suita-city, Osaka, Japan. A summary

TABLE I
BASIC PERSONAL INFORMATION OF CONTRIBUTORS. x, y, AND z IN x-y-z MEAN THE MINIMUM, AVERAGE, AND MAXIMUM VALUES, RESPECTIVELY

Item	Group A	Group B
Age	20-27-48	21-27-62
Height	147-166-185	148-167-185
Weight	43-56-75	43-59-88
Household income	0-529-2000	0-527-2000
Personal annual income	0-89-650	0-121-1050
Monthly living expenses	2-28-300	2-16-90

TABLE II
DESCRIPTION OF VIDEO CONTENTS

Item	Group A	Group B
No. of short videos	240	153
Length of short videos	15 sec	14–24 sec
No. of long videos	120	161
Length of long videos	30 sec	25–44 sec
Broadcasting region	Japan	Japan
Broadcasting period	July 2011 –Dec. 2017	July 2015 –March 2017

of the basic personal information is provided in Table I. Further personal information of the contributors can be found in our previous paper [3]. The numbers of contributors in groups A and B are large enough compared with the ones in prior works [22], [24], [30], [31]; typically, the numbers have ranged from 20 to 40. During the experiment, each contributor watched a mixed sequence of short and long videos (Table II) for approximately two hours, which is also sufficiently long compared with prior works [30], [34]–[36].

A. Brain Information and Model

We used the brain information obtained in our previous work [3]. The information pertaining to the sets of video content used for contributor groups A and B is summarized in Table II. These included actual TV broadcasts and online advertising materials. We conducted experiments using fMRI to acquire the brain information from the two contributor groups. The experimental protocol was approved by the ethics and safety committees of the National Institute of Information and Communications Technology (NICT). To acquire the movie stimuli, a set of video contents was sequentially concatenated in a pseudo-random order, after which 14 nonoverlapping video sequences were obtained. The individual sequences were displayed in separate scans. Twelve of the sequences were presented once each, while the other two were presented four times each in four separate scans. The fMRI responses to these two sequences were averaged across the four scans. Further information about the experiment is available in our previous paper [3].

The brain model, Δ, was obtained by using the method presented in our previous paper [3]. We utilized principal component analysis (PCA) to reduce the number of dimensions of the brain information signals from the original number of dimensions, which was around 60 000.[1] Dynamic time warping (DTW) was used to quantify the similarity of the brain information between each pair of video contents. That is, a matrix that represents the similarity between each pair of video contents as the distance between them was obtained for each contributor. Then, the distance between the matrices of each pair of contributors was measured by means of the distance between two network graphs in the weighted graph-matching problem (WGMP). Δ is a matrix where the elements δ_{ij} are the distances between each pair of contributors measured by the WGMP.

[1]Group A: mean 62 282, standard deviation 5198. Group B: mean 62 211, standard deviation 5243.

B. Profile Information and Model

We created a list of questionnaire items, P, to obtain the profile information from contributors. The categories and subcategories of these questionnaire items are listed in Table III. The total number of items was 209. However, we removed ones not suitable for this work, such as those with a free answer format. We then obtained a list of the answers of each contributor for each questionnaire item, A_v. As discussed in Section III-D, the profile model, D, is a matrix that represents the differences of the profile information between each pair of contributors. The element of D, d_{ij}, is given by calculating the absolute difference of the answers between contributors i and j. That is, d_{ij} is a vector. To obtain this vector type of data from categorical data types of answers, we simply counted the difference between two answers neighboring each other as one (if there is a "two hop" between two answers, the distance would simply be two).

C. Machine Learning and Feature Selection

As we explained in Section III-D, the ML model created by the model converter has the matrices of the profile and brain models as its input and output, respectively. We created the ML model by using Δ and D obtained by the means described in Sections IV-A and IV-B. The input variable of ML is d_{ij}, which is one of the elements of D and indicates the difference of the profile model between contributors i and j as a vector. The output variable of ML is δ_{ij}, which is one of the elements of Δ and indicates the difference between contributors i and j as a real number. That is, actually obtained data for the sets of d_{ij} and δ_{ij} are used for training and validation in ML. We used random forest (RF) [37] as the ML method for the following reasons. First, RF can perform regression tasks effectively while setting a smaller number of parameters than more complicated ML methods, such as deep learning [38]. Second, unlike other simple regression methods, such as support vector machine (SVM) and linear regression, RF provides the impurity metric, which enables us to perform feature selection without additional computation. When we used RF in the evaluation of this study, 80% and 20% of the data were used for training and validation, respectively. The number of trees was set to ten. Mean square error (MSE) was used as the criterion for the decision tree. We utilized Scikit-learn, which is a well-known and widely used toolset for ML.

We used two feature selection methods. The first one is the impurity method [32], which is applicable for RF but not for neural network-based ML methods. In RF, at each node of the decision tree, m elements are selected at random out of

TABLE III
CATEGORIES AND SUBCATEGORIES OF QUESTIONNAIRE ITEMS

Category	Subcategory
Basic attributes	Sex
	Age
	Age (range)
	Zodiac sign
	Blood type
	Academic background
	University (major)
	Dominant hand
	Dominant foot
	Digit ratio
	Birthplace
	Height and weight
	Residence
	Family
	Annual income
Education and development	Six factors of parenting
	Clusters of parenting
Job	Profession
	Place of work
Private life	Lifestyle
	Hobbies
	Key life events
	Cost of living
Physical health	Sleep
	Exercise habits
	Medical history
Mental health	Subjective well-being
	Four factors of well-being
Mental illness	Gambling addiction
	Nicotine addiction
	Alcohol addiction
Personality	Big Five (TIPI-J)
	Self-control
Behavioral economy	Decision making
	Risk aversion
	Bubble reaction
	Rational decision making
	Over-confidence
Culture	Schwartz's theory of basic values

the total number of features, and the best split among them is selected. At each node t within the binary trees T of RF, the optimal split is sought using the impurity $i(t)$, which is a computationally efficient approximation of the entropy, measuring how well a potential split separates the samples of the two groups in this particular node. This means that the impurity reflects the importance of the element used for splitting. The second one is the permutation method [33], which is a model-inspection method that can be used for any fitted estimator when the data are rectangular. This is especially useful for nonlinear or opaque estimators. The importance extracted by permutation is defined as the decrease in a model score when a single feature value is randomly shuffled. This procedure breaks the relationship between the feature and target; a drop in the model score is indicative of how much the prediction accuracy depends on the feature. This method benefits from being model-agnostic and can be performed many times with different permutations of the feature. The purpose of using two methods for feature selection is to show that the superiority of the proposed scheme does not depend on only a specific feature selection method, rather than to compare which method is better.

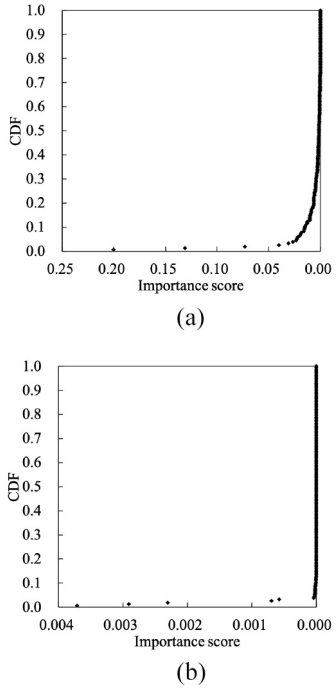


Fig. 3. CDF of importance scores. Number of PCA dimensions: 1000. Contributor group: B (no. of contributors: 39). Video length: short. (a) Impurity. (b) Permutation.

It should be noted that the proposed scheme does not limit the ML methods that can be used in it; basically, any existing ML method that work for regression tasks is applicable to the proposed scheme. If other ML methods (e.g., SVM, linear regression, or deep learning) are applied to the proposed scheme, the impurity method, which is the feature selection method specific to RF, becomes inapplicable. In contrast, the permutation method, which is a versatile feature selection method, is still applicable even after the base ML method changes.

It is also important to note that the proposed scheme can use any type of feature selection method that has previously been proposed [39]–[41]. However, the present study does not aim to determine the best combination of ML and feature selection methods for our proposed scheme but rather to demonstrate the validity of our proposed scheme using standard ML and feature selection methods. For this purpose, the impurity and permutation methods with RF would be sufficient for evaluating the proposed scheme's effectiveness, namely, how consistent it is across different types of feature selection methods.

V. RESULTS

A. Distribution of Importance Scores

Fig. 3(a) and (b) plots the cumulative distribution functions (CDFs) of the importance scores of each questionnaire item obtained using the impurity and permutation methods, respectively. As seen in Fig. 3(a), only around 20% of the top questionnaire items had high importance scores, while the rest of them had low ones. This suggests that using only the top 20% of questionnaire items is satisfactory for inferring

TABLE IV
RESULTS OBTAINED USING IMPURITY METHOD. G., L., AND D. MEAN CONTRIBUTOR GROUP, VIDEO LENGTH, AND NUMBER OF PCA DIMENSIONS, RESPECTIVELY

G.	A (No. of contributors: 26)				B (No. of contributors: 39)			
L.	Short		Long		Short		Long	
D.	1000	2000	1000	2000	1000	2000	1000	2000
10% score	1.8×10^{-2}	2.0×10^{-2}	1.7×10^{-2}	1.8×10^{-2}	1.5×10^{-2}	1.4×10^{-2}	1.6×10^{-2}	1.5×10^{-2}
15 % score	1.3×10^{-2}	1.4×10^{-2}	1.3×10^{-2}	1.3×10^{-2}	9.6×10^{-3}	8.9×10^{-3}	1.3×10^{-2}	1.2×10^{-2}
20% score	9.0×10^{-3}	1.0×10^{-2}	1.0×10^{-2}	9.7×10^{-3}	6.3×10^{-3}	6.6×10^{-3}	9.6×10^{-3}	8.4×10^{-3}
50% score	1.1×10^{-3}	1.0×10^{-3}	9.8×10^{-4}	7.5×10^{-4}	1.4×10^{-3}	1.3×10^{-3}	1.6×10^{-3}	1.5×10^{-3}
R^2	0.37	0.50	0.53	0.38	0.64	0.63	0.53	0.47

TABLE V
RESULTS OBTAINED USING PERMUTATION METHOD. G., L., AND D. MEAN CONTRIBUTOR GROUP, VIDEO LENGTH, AND NUMBER OF PCA DIMENSIONS, RESPECTIVELY

G.	A (No. of contributors: 26)				B (No. of contributors: 39)			
L.	Short		Long		Short		Long	
D.	1000	2000	1000	2000	1000	2000	1000	2000
10% score	1.2×10^{-5}	1.7×10^{-5}	9.6×10^{-6}	1.2×10^{-5}	9.9×10^{-6}	1.2×10^{-5}	1.3×10^{-5}	1.5×10^{-5}
15 % score	6.6×10^{-6}	4.3×10^{-6}	6.2×10^{-6}	4.5×10^{-6}	3.0×10^{-6}	5.1×10^{-6}	6.2×10^{-6}	8.5×10^{-6}
20% score	2.7×10^{-6}	2.8×10^{-6}	3.5×10^{-6}	1.8×10^{-6}	1.9×10^{-6}	3.1×10^{-6}	3.9×10^{-6}	4.6×10^{-6}
50% score	4.2×10^{-8}	3.1×10^{-8}	2.4×10^{-8}	2.7×10^{-8}	4.7×10^{-8}	7.1×10^{-8}	8.6×10^{-8}	1.2×10^{-7}
R^2	0.35	0.34	0.21	0.36	0.60	0.65	0.63	0.58

the brain model from the profile model. Next, looking at Fig. 3(b), only around 10% of the top questionnaire items had high importance scores. This also suggests that in the case of using the permutation method, only the top 10% of questionnaire items are necessary for the inferring.

Table IV lists the top 10, 15, 20, and 50% importance scores in the case of using the impurity method. We compared the importance scores for various combinations of parameters for contributor groups, video lengths, and the number of PCA dimensions. In the table, in all case, the top 10% of importance score is ten times or more than the top 50% of importance score. This is consistent with the results shown in Fig. 3(a), where only the top ranked questionnaire items had high importance scores. Table IV also shows the R^2 coefficient of determination. In most cases, the coefficients were larger than or equal to 0.5, though in three cases they were a little lower than 0.5.

Table V lists the importance scores in the case of using the permutation method. Here, as in Table IV, the 10% importance scores were much higher than the 50% ones, though the difference was 100 times or more. Regarding the R^2 coefficient of determination, the coefficients were not so large for group A unfortunately, but they were satisfactory for group B.

B. Identification of Top Ranked Questionnaire Items

This section compares the top ranked questionnaire items obtained using different parameters or feature selection methods. From the comparison, we will verify that the proposed scheme can identify the top questionnaire items that contribute most significantly to the inference of brain models.

Table VI lists the ratio of common items and Spearman's rank correlation coefficients [42] between two lists of the top 20% or 10% ranked questionnaire items obtained with two different numbers of PCA dimensions, 1000 and 2000. The ratio of common items equals the number of common items divided by the total number of items in the list, which was 31 and 15 in the cases of the top 20% and 10%, respectively.

TABLE VI
RATIO OF COMMON ITEMS AND SPEARMAN'S RANK CORRELATION COEFFICIENT BETWEEN TWO LISTS OF TOP 20% OR 10% RANKED QUESTIONNAIRE ITEMS OBTAINED WITH TWO DIFFERENT NUMBERS OF PCA DIMENSIONS (1000 AND 2000). G. AND L. MEAN CONTRIBUTOR GROUP AND VIDEO LENGTH, RESPECTIVELY. IMPUR. AND PERMUT. MEAN THE IMPURITY AND PERMUTATION METHODS, RESPECTIVELY

Impur.		Top 20%		Top 10%	
G.	L.	Common	Spearman's	Common	Spearman's
A	Short	0.58	0.41	0.53	0.62
A	Long	0.71	0.40	0.47	0.43
B	Short	0.68	0.66	0.73	0.87
B	Long	0.81	0.65	0.67	0.76

Permut.		Top 20%		Top 10%	
G.	L.	Common	Spearman's	Common	Spearman's
A	Short	0.58	0.41	0.67	0.60
A	Long	0.71	0.40	0.53	0.45
B	Short	0.77	0.73	0.60	0.83
B	Long	0.68	0.90	0.87	0.92

We extracted the common items from one of the two lists and generated a new list that includes only the common items and did the same process for the other list. Then, we calculated Spearman's rank correlation coefficient between the two newly generated lists to evaluate the similarity of their rank orders. As shown in Table VI, the ratio of common items was around 0.5 to 0.9, which is quite high. This means that similar questionnaire items were ranked with high importance in the two compared lists even though their PCA dimensions were different. Spearman's rank correlation coefficients were larger than 0.4, which means that common items between them were ranked similarly. In particular, in the top 10% list for group B, long length of video, and the permutation method, the ratio of common items and Spearman's rank correlation coefficient were 0.87 and 0.92, which is quite successful.

Table VII shows the ratio of common items and Spearman's rank correlation coefficient between two lists obtained with short and long videos. The ratios of common items were around 0.3 to 0.7, which were smaller than the ones in

TABLE VII
RATIO OF COMMON ITEMS AND SPEARMAN'S RANK CORRELATION COEFFICIENT BETWEEN TWO LISTS OF TOP 20% OR 10% RANKED QUESTIONNAIRE ITEMS OBTAINED WITH SHORT AND LONG VIDEOS. G. AND D. MEAN CONTRIBUTOR GROUP AND NUMBER OF PCA DIMENSIONS, RESPECTIVELY. IMPUR. AND PERMUT. MEAN THE IMPURITY AND PERMUTATION METHODS, RESPECTIVELY

Impur.		Top 20%		Top 10%	
G.	D.	Common	Spearman's	Common	Spearman's
A	1000	0.61	0.29	0.33	0.00
A	2000	0.61	0.15	0.40	0.60
B	1000	0.74	0.44	0.60	0.63
B	2000	0.77	0.39	0.53	0.90

Permut.		Top 20%		Top 10%	
G.	D.	Common	Spearman's	Common	Spearman's
A	1000	0.61	0.36	0.40	0.29
A	2000	0.61	0.84	0.47	0.18
B	1000	0.68	0.70	0.60	0.70
B	2000	0.65	0.60	0.47	0.54

TABLE VIII
RATIO OF COMMON ITEMS AND SPEARMAN'S RANK CORRELATION COEFFICIENT BETWEEN TWO LISTS OF TOP 20% OR 10% RANKED QUESTIONNAIRE ITEMS OBTAINED WITH DIFFERENT FEATURE SELECTION METHODS, IMPURITY AND PERMUTATION. L. AND D. MEAN VIDEO LENGTH AND NUMBER OF PCA DIMENSIONS, RESPECTIVELY. GROUPS A AND B MEAN CONTRIBUTOR GROUPS A AND B, WHICH INCLUDED 27 AND 40 CONTRIBUTORS, RESPECTIVELY

Group A		Top 20%		Top 10%	
L.	D.	Common	Spearman's	Common	Spearman's
Short	1000	0.68	0.44	0.53	0.64
Short	2000	0.68	0.33	0.60	−0.30
Long	1000	0.74	0.44	0.60	0.73
Long	2000	0.61	0.35	0.47	0.50

Group B		Top 20%		Top 10%	
L.	D.	Common	Spearman's	Common	Spearman's
Short	1000	0.68	0.65	0.67	0.84
Short	2000	0.71	0.85	0.67	0.95
Long	1000	0.71	0.67	0.60	0.78
Long	2000	0.74	0.75	0.73	0.75

Table VI. Spearman's rank correlation coefficients in Table VII were also smaller than the ones in Table VI. Particularly, in the top 10% list for group A, 1000 dimensions of PCA, and the impurity method, Spearman's rank correlation coefficient was around 0.00. This is because, as shown in Table IV, the R^2 coefficient of determination for the impurity method, group A, short video, and 1000 dimensions of PCA was small. This suggests that the R^2 coefficient of determination works well as a metric to ensure the accuracy of top ranked questionnaire items by using a feature selection method.

Table VIII presents the ratio of common items and Spearman's rank correlation coefficient between two lists obtained using different feature selection methods, impurity and permutation. As shown, the overall performance was good: the ratios of common items were around 0.5 to 0.8, while Spearman's rank correlation coefficient was around 0.3 to 1.0. The only exception was a negative correlation in the top 10% list for group A, 2000 dimensions of PCA, and short video. As discussed above, this is because, as shown in Table V, the R^2 coefficient of determination for the permutation method, contributor group A, short video, and 2000 dimensions of PCA was small. The R^2 coefficient of determination works as a metric to ensure the accuracy of top ranked questionnaire items.

TABLE IX
EXAMPLES OF TOP RANKED ITEMS. COMPARISON BETWEEN PCA DIMENSIONS. FEATURE SELECTION METHOD: PERMUTATION. CONTRIBUTOR GROUP: B. VIDEO LENGTH: LONG

Rank	No. of PCA dimensions: 1000
1	hobbies-walking and eating
2	over-confidence-stock price prediction
3	personal annual income
4	hobbies-reading comic books
5	hobbies-motorbikes
6	hobbies-studying
7	dominant hand
8	dominant foot
9	decision making-loss aversion (λ)
10	over-confidence-other
11	six factors of parenting-uninterested
12	residence-prefectures
13	six factors of parenting-independent
14	four factors of well-being-thanks
15	hobbies-none
16	six factors of parenting-scolding experience
17	birthplace
18	hobbies-cooking
19	household annual income
20	past medical history-respiratory disease
21	bubble reaction-number of detections
22	over-confidence-investment
23	height and weight-weight
24	subjective well-being
25	digit ratio
26	sleep debt
27	Schwartz's theory of basic values-benevolence
28	Big Five-TiPIJ-conscientiousness
29	bubble reaction-profit and loss
30	risk aversion-total time
31	cost of living-monthly cost of living

Rank	No. of PCA dimensions: 2000
1	over-confidence-stock price prediction
2	hobbies-walking and eating
3	personal annual income
4	hobbies-comic books
5	hobbies-motorbikes
6	dominant hand
7	over-confidence-other
8	dominant foot
9	decision making-loss aversion (λ)
10	hobbies-studying
11	hobbies-gambling
12	six factors of parenting-scolding experience
13	six factors of parenting-independent
14	residence-prefectures
15	four factors of well-being-gratitude
16	life event-divorce
17	over-confidence-investment
18	height and weight-weight
19	height and weight-height
20	hobbies-none
21	six factors of parenting-shared time
22	zodiac sign
23	risk aversion-total time
24	age
25	six factors of parenting-uninterested
26	bubble reaction-number of detections
27	bubble reaction-profit and loss
28	Schwartz's theory of basic values-security
29	Profession
30	Big Five-TiPIJ-neuroticism
31	sex

TABLE X
EXAMPLES OF TOP RANKED ITEMS. COMPARISON BETWEEN FEATURE SELECTION METHODS. NUMBER OF PCA DIMENSIONS: 2000. CONTRIBUTOR GROUP: B. VIDEO LENGTH: SHORT

Rank	Permutation
1	hobbies-motorbikes
2	hobbies-walking and eating
3	over-confidence-stock price prediction
4	personal annual income
5	decision making-loss aversion (λ)
6	sleep debt
7	hobbies-reading comic books
8	hobbies-gambling
9	life event-divorce
10	over-confidence-investment
11	sex
12	hobbies-none
13	decision making-probability weighting (σ)
14	height and weight-weight
15	birthplace
16	dominant hand
17	bubble reaction-profit and loss
18	past medical history-none
19	dominant foot
20	residence-prefectures
21	past medical history-allergies
22	self-control
23	six factors of parenting-scolding experience
24	over-confidence-other
25	risk aversion-total time
26	Schwartz's theory of basic values-benevolence
27	alcohol addiction
28	household annual income
29	decision making-value (α)
30	six factors of parenting-uninterested
31	Big Five-TiPIJ-agreeableness

Rank	Impurity
1	hobbies-motorbikes
2	over-confidence-stock price prediction
3	hobbies-walking and eating
4	personal annual income
5	decision making-loss aversion (λ)
6	sleep debt
7	hobbies-gambling
8	over-confidence-other
9	four factors of well-being-thanks
10	six factors of parenting-set an example
11	birthplace
12	height and weight-height
13	decision making-probability weighting (σ)
14	six factors of parenting-independent
15	height and weight-weight
16	bubble reaction-profit and loss
17	over-confidence-investment
18	life event-divorce
19	subjective well-being
20	household annual income
21	four factors of well-being-let's try
22	cost of living-monthly cost of living
23	dominant hand
24	six factors of parenting-uninterested
25	six factors of parenting-shared time
26	six factors of parenting-scolding experience
27	dominant foot
28	alcohol addiction
29	self-control
30	decision making-value (α)
31	Big Five-TiPIJ-conscientiousness

Considering the results shown above, we highlight two results with a large ratio of common items and large Spearman's coefficient from Tables VI–VIII and show the top ranked questionnaire items in the lists, which are shown in Tables IX and X. Table IX shows the top ranked questionnaire items in the lists of two PCA dimensions, 1000 or 2000, with the permutation method, contributor group B, and long video. As shown, common items were listed in a similar order in both lists. Table X shows the top ranked questionnaire items in the lists of two feature selection methods, impurity and permutation, with 2000 dimensions of PCA, contributor group B, and short video. As shown, common items were listed in a similar order in both lists.

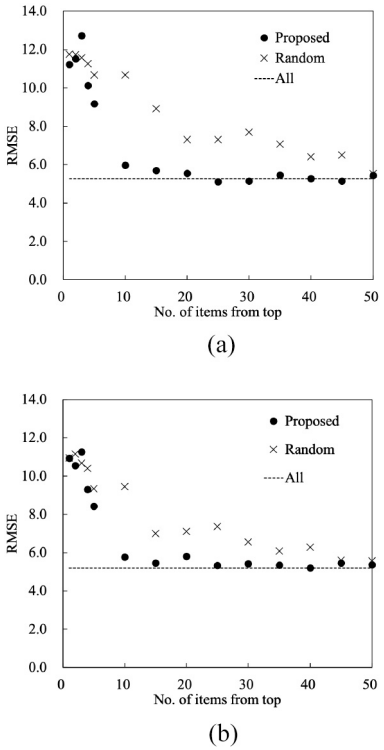


Fig. 4. RMSE versus the number of items from top. Feature selection method: permutation. Contributor group: B (no. of contributors: 39). Video length: long. (a) Number of PCA dimensions: 1000. (b) Number of PCA dimensions: 2000.

C. Evaluation of Inference Accuracy Using Top Ranked Items

Finally, we evaluate how the inference accuracy varies with using top ranked items. To evaluate the inference accuracy of the brain model from the profile model, we picked and masked one of the differences between each pair of users in the brain model, which corresponds to δ_{ij} in Section IV-C, and estimated it from other values of the differences between each pair of users in the brain model. Then, we measured the root MSE (RMSE) of the estimated value against the masked one. Figs. 4 and 5 plot RMSE versus the number of items from the top. The random scheme in the figure just sorts items randomly without considering the importance. All case means using all of the 209 items. The parameters for Figs. 4 and 5 correspond to the ones in Tables IX and X.

In Figs. 4 and 5, for the proposed scheme and the random scheme, RMSEs decreased as the number of items from the top increased, which is expected because a greater variety of

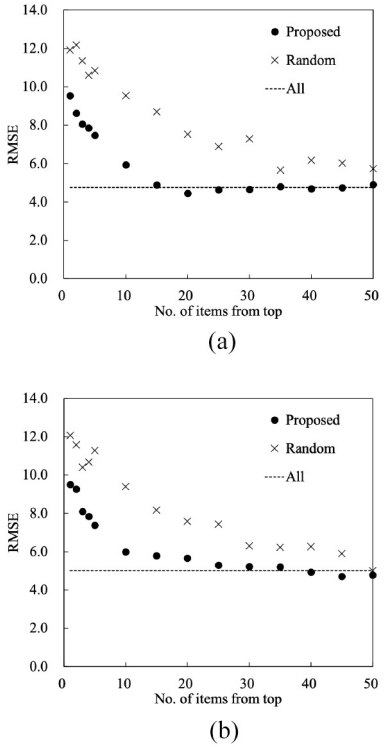


Fig. 5. RMSE versus the number of items from top. Number of PCA dimensions: 2000. Contributor group: B (number of contributors: 39). Video length: short. (a) Feature selection method: permutation. (b) Feature selection method: impurity.

items improves the inference accuracy. The proposed scheme achieved almost the same level of inference accuracy as the all case by only using around 15 to 20 items from the top, which is consistent with the distribution of importance scores discussed in Section V-A; only top 10% items have high importance scores. However, the random scheme required 50 items. Thus, the above observation suggests that the proposed scheme can reduce 209 items to just 15 to 20 items while maintaining the inference accuracy, which leads to the reduction of the cost of personal information collection for users.

VI. Conclusion

In this article, we have proposed a scheme that resolves the tradeoff between the inference performance of a brain model from a profile model and the cost of collecting profile information. In the proposed scheme, we use an ML model to infer the brain model from the profile model and use a feature selection method to reduce the cost of collecting profile information, i.e., the number of questionnaire items. By using only the top questionnaire items with the highest importance scores, the proposed scheme maintains the inference performance as high as possible while limiting the number of questionnaire items. We demonstrated the effectiveness of the proposed scheme with a performance evaluation using an experimentally obtained brain model and a profile model created from actual profile information. The results over different experimental parameters, video lengths, and feature selection methods demonstrate that the proposed scheme successfully identifies the top questionnaire items that contribute most significantly to the inference of brain models when the ML model for the inference is created accurately, which is ensured by a metric, such as the R^2 coefficient of determination. Finally, we presented the performance evaluation of the inference accuracy. The results suggest that the proposed scheme can reduce the number of questionnaire items while maintaining the inference accuracy, which leads to the reduction of the cost of personal information collection for users.

An important direction for future investigation will be to determine the best combination of ML and feature selection methods with respect to the performance optimization of the proposed scheme. Since the proposed scheme accepts any type of ML or feature selection method, as we mentioned in Section IV-C, we need to test various sophisticated ML and feature selection methods that have been previously proposed [39]–[41]. At the same time, we will also need to reduce the total computation cost of the ML and feature selection methods for real-world applications, although this problem was not formulated in this article. In this respect, the combination of RF and the impurity method we used in this study is highly effective because, as we mentioned in Section IV-C, RF is a simple ML method with a small number of parameters, and the impurity method does not require additional computation for feature selection. Nevertheless, further analyses through numerical evaluation will be needed to determine which combination of ML and feature selection methods is most effective in terms of computation cost.

Acknowledgment

The authors are grateful to Fumihiro Takayama and Yoshifumi Aimoto of NTT DATA Institute of Management Consulting, Inc. for sharing datasets from "Human info DB" for this work. They also thank Yuichi Inagaki, Shuichirou Shimizu, and Naoya Kawamura, who are students at Kyoto University, and Keiichiro Sato, who graduated from Kyoto University in March 2020, for their assistance.

References

[1] C. S. Soon, M. Brass, H.-J. Heinze, and J.-D. Haynes, "Unconscious determinants of free decisions in the human brain," *Nature Neurosci.*, vol. 11, no. 5, pp. 543–545, 2008.

[2] A. Tusche, S. Bode, and J.-D. Haynes, "Neural responses to unattended products predict later consumer choices," *J. Neurosci.*, vol. 30, no. 23, pp. 8024–8031, 2010.

[3] R. Shinkuma, S. Nishida, M. Kado, N. Maeda, and S. Nishimoto, "Relational network of people constructed on the basis of similarity of brain activities," *IEEE Access*, vol. 7, pp. 110258–110266, 2019.

[4] S. Nishida and S. Nishimoto, "Decoding naturalistic experiences from human brain activity via distributed representations of words," *NeuroImage*, vol. 180, pp. 232–242, Oct. 2018.

[5] J. P. Dmochowski, M. A. Bezdek, B. P. Abelson, J. S. Johnson, E. H. Schumacher, and L. C. Parra, "Audience preferences are predicted by temporal reliability of neural processing," *Nature Commun.*, vol. 5, no. 1, pp. 1–9, 2014.

[6] V. Venkatraman *et al.*, "Predicting advertising success beyond traditional measures: New insights from neurophysiological methods and market response modeling," *J. Mark. Res.*, vol. 52, no. 4, pp. 436–452, 2015.

[7] S. Nishida, Y. Nakano, A. Blanc, N. Maeda, M. Kado, and S. Nishimoto, "Brain-mediated transfer learning of convolutional neural networks," in *Proc. 34th AAAI Conf. Artif. Intell.*, 2020, pp. 5281–5288.

[8] L. C. Tong, M. Y. Acikalin, A. Genevsky, B. Shiv, and B. Knutson, "Brain activity forecasts video engagement in an Internet attention market," *Proc. Nat. Acad. Sci. USA*, vol. 117, no. 12, pp. 6936–6941, 2020.
[9] P. Papadopoulos, N. Kourtellis, and E. P. Markatos, "The cost of digital advertisement: Comparing user and advertiser views," in *Proc. World Wide Web Conf.*, 2018, pp. 1479–1489.
[10] J. Tang, T. Lou, and J. Kleinberg, "Inferring social ties across heterogenous networks," in *Proc. 5th ACM Int. Conf. Web Search Data Mining*, 2012, pp. 743–752.
[11] M. Fang, J. Yin, and X. Zhu, "Transfer learning across networks for collective classification," in *Proc. IEEE 13th Int. Conf. Data Mining*, 2013, pp. 161–170.
[12] Y. Wang, C. Kang, L. M. Bettencourt, Y. Liu, and C. Andris, "Linked activity spaces: Embedding social networks in urban space," in *Computational Approaches for Urban Environments*. Cham, Switzerland: Springer, 2015, pp. 313–336.
[13] L. Wang, H. Wu, W. Wang, and K.-C. Chen, "Socially enabled wireless networks: Resource allocation via bipartite graph matching," *IEEE Commun. Mag.*, vol. 53, no. 10, pp. 128–135, Oct. 2015.
[14] R. Shinkuma, Y. Sawada, Y. Omori, K. Yamaguchi, H. Kasai, and T. Takahashi, "A socialized system for enabling the extraction of potential values from natural and social sensing," in *Modeling and Processing for Next-Generation Big-Data Technologies*. Cham, Switzerland: Springer, 2015, pp. 385–404.
[15] Z. Ding, Y. Omori, R. Shinkuma, and T. Takahashi, "Designing mobility models based on relational graph," *IEICE Trans. Inf. Syst.*, vol. 97, no. 12, pp. 3007–3015, 2014.
[16] Z. Ding, R. Shinkuma, and T. Takahashi, "Designing temporally and spatially integrated social mobility models for wireless network researches," in *Proc. 15th Asia-Pac. Netw. Oper. Manage. Symp. (APNOMS)*, 2013, pp. 1–5.
[17] R. Shinkuma *et al.*, "Toward future network systems boosting interactions between people in social networks," in *Proc. 20th Int. Conf. Comput. Commun. Netw. (ICCCN)*, 2011, pp. 1–6.
[18] W. Fan *et al.*, "Graph neural networks for social recommendation," in *Proc. World Wide Web Conf.*, 2019, pp. 417–426.
[19] Q. Dai, X. Shen, X.-M. Wu, and D. Wang, "Network transfer learning via adversarial domain adaptation with graph convolution," 2019. [Online]. Available: arXiv:1909.01541.
[20] X. Shen, Q. Dai, F.-L. Chung, W. Lu, and K.-S. Choi, "Adversarial deep network embedding for cross-network node classification," 2020. [Online]. Available: arXiv:2002.07366.
[21] J. Dubois and R. Adolphs, "Building a science of individual differences from fMRI," *Trends Cogn. Sci.*, vol. 20, no. 6, pp. 425–443, 2016.
[22] M. P. Van Den Heuvel, C. J. Stam, R. S. Kahn, and H. E. H. Pol, "Efficiency of functional brain networks and intellectual performance," *J. Neurosci.*, vol. 29, no. 23, pp. 7619–7624, 2009.
[23] E. S. Finn *et al.*, "Functional connectome fingerprinting: Identifying individuals using patterns of brain connectivity," *Nature Neurosci.*, vol. 18, no. 11, pp. 1664–1671, 2015.
[24] J. S. Adelstein *et al.*, "Personality is reflected in the brain's intrinsic functional architecture," *PloS One*, vol. 6, no. 11, 2011, Art. no. e27633.
[25] R. E. Beaty *et al.*, "Robust prediction of individual creative ability from brain functional connectivity," *Proc. Nat. Acad. Sci. USA*, vol. 115, no. 5, pp. 1087–1092, 2018.
[26] R. L. Buckner, F. M. Krienen, and B. T. T. Yeo, "Opportunities and limitations of intrinsic functional connectivity MRI," *Nature Neurosci.*, vol. 16, no. 7, pp. 832–837, 2013.
[27] E. S. Finn, D. Scheinost, D. M. Finn, X. Shen, X. Papademetris, and R. T. Constable, "Can brain state be manipulated to emphasize individual differences in functional connectivity?" *Neuroimage*, vol. 160, pp. 140–151, Oct. 2017.
[28] J. Gonzalez-Castillo and P. A. Bandettini, "Task-based dynamic functional connectivity: Recent findings and open questions," *Neuroimage*, vol. 180, pp. 526–533, Oct. 2018.
[29] A. S. Greene, S. Gao, D. Scheinost, and R. T. Constable, "Task-induced brain state manipulation improves prediction of individual traits," *Nature Commun.*, vol. 9, no. 1, pp. 1–13, 2018.
[30] C. Parkinson, A. M. Kleinbaum, and T. Wheatley, "Similar neural responses predict friendship," *Nature Commun.*, vol. 9, no. 1, pp. 1–14, 2018.
[31] I. Charest, R. A. Kievit, T. W. Schmitz, D. Deca, and N. Kriegeskorte, "Unique semantic space in the brain of each beholder predicts perceived similarity," *Proc. Nat. Acad. Sci. USA*, vol. 111, no. 40, pp. 14565–14570, 2014.
[32] R. Shinkuma and T. Nishio, "Data assessment and prioritization in mobile networks for real-time prediction of spatial information with machine learning," in *Proc. IEEE 1st Int. Workshop Netw. Meets Intell. Comput. (NMIC)*, 2019, pp. 1–6.
[33] (2019). *Scikit Learn, 4.2. Permutation Feature Importance*. Accessed: Apr. 19, 2020. [Online]. Available: https://scikit-learn.org/stable/modules/permutation_importance.html
[34] U. Hasson, Y. Nir, I. Levy, G. Fuhrmann, and R. Malach, "Intersubject synchronization of cortical activity during natural vision," *Science*, vol. 303, no. 5664, pp. 1634–1640, 2004.
[35] A. G. Huth, S. Nishimoto, A. T. Vu, and J. L. Gallant, "A continuous semantic space describes the representation of thousands of object and action categories across the human brain," *Neuron*, vol. 76, no. 6, pp. 1210–1224, 2012.
[36] J. V. Haxby *et al.*, "A common, high-dimensional model of the representational space in human ventral temporal cortex," *Neuron*, vol. 72, no. 2, pp. 404–416, 2011.
[37] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[38] S. Raschka, *Python Machine Learning*. Birmingham, U.K.: Packt Publ. Ltd, 2015.
[39] B. Shen, W. Xie, and Z. J. Kong, "Clustered discriminant regression for high-dimensional data feature extraction and its applications in healthcare and additive manufacturing," *IEEE Trans. Autom. Sci. Eng.*, early access, Oct. 22, 2020, doi: 10.1109/TASE.2020.3029028.
[40] M. Ghahramani, Y. Qiao, M. C. Zhou, A. O. Hagan, and J. Sweeney, "Ai-based modeling and data-driven evaluation for smart manufacturing processes," *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 4, pp. 1026–1037, Jul. 2020.
[41] H. Liu, M. Zhou, and Q. Liu, "An embedded feature selection method for imbalanced data classification," *IEEE/CAA J. Automatica Sinica*, vol. 6, no. 3, pp. 703–715, May 2019.
[42] T. D. Gautheir, "Detecting trends using Spearman's rank correlation coefficient," *Environ. Forensics*, vol. 2, no. 4, pp. 359–362, 2001.

Ryoichi Shinkuma (Senior Member, IEEE) received the B.E., M.E., and Ph.D. degrees in communications engineering from Osaka University, Suita, Japan, in 2000, 2001, and 2003, respectively.

He joined the Graduate School of Informatics, Kyoto University, Kyoto, Japan, and worked there as an Assistant Professor from 2003 to 2011 and as an Associate Professor from 2011 to 2021. He was a Visiting Scholar with the Wireless Information Network Laboratory, Rutgers University, New Brunswick, NJ, USA, from Fall 2008 to Fall 2009. In 2021, he joined the Faculty of Engineering, Shibaura Institute of Technology, Tokyo, Japan, as a Professor. His main research interest is cooperation in heterogeneous networks.

Prof. Shinkuma received the Young Researchers' Award from the IEICE in 2006 and the Young Scientist Award from Ericsson Japan in 2007, the TELECOM System Technology Award from the Telecommunications Advancement Foundation in 2016 and the Best Tutorial Paper Award from the IEICE Communications Society in 2019. He was the Chairperson of the Mobile Network and Applications Technical Committee of the IEICE Communications Society from 2017 to 2019. He is a Fellow of the IEICE.

Satoshi Nishida received the Ph.D. degree in medicine from Kyoto University, Kyoto, Japan, in 2014.

After short-term postdoctoral work with Kyoto University, he joined Center for Information and Neural Networks, National Institute of Information and Communications Technology, Suita, Japan, in 2014, where he is currently a Senior Researcher. As of 2020, he is also a Guest Associate Professor with the Graduate School of Frontier Biosciences, Osaka University, Suita. His current research interests include the areas of cognitive neuroscience, experimental psychology, and artificial intelligence.

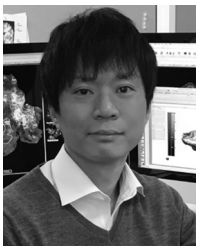
Dr. Nishida received the Young Researcher Award from the IEEE CIS Japan Chapter in 2011 and the Best Research Award from the Japan Neural Network Society in 2011.

Naoya Maeda graduated from Department of Electrical and Electronic Engineering, Faculty of Engineering, Kanto Gakuin University, Yokohama, Japan. He joined the Software House as a new graduate in 1997, where he was engaged in system development for the telecom, engineering, entertainment, medical, and research industries.

He joined NTT DATA, Tokyo, Japan, in 2005, where he was engaged in project management for corporate system development and business support by customer residency. Since 2014, he has been working on applied research and commercialization of integrated neuroscience and artificial intelligence technology through industry-academia collaborations.

Masataka Kado received the Master degree in computer science, from the Graduate School of Engineering, Nihon University, Tokyo, Japan, in 2003.

He has been working with NTT DATA Corporation, Tokyo, Japan, since 2003. He was engaged in research and development on network security and the cloud from 2003 to 2010 and engaged in applied research on advanced technologies relating to neuroscience and human sensing from 2011 to 2015. Since 2016, he has been working on the planning and commercialization of neuro-business.



Shinji Nishimoto received the Ph.D. degree in neurophysiology from Osaka University, Suita, Japan, in 2005.

He worked as a Postdoctoral Fellow and as an Associate Specialist with Helen Wills Neuroscience Institute, University of California, Berkeley, CA, USA, from 2005 to 2013. He joined CiNet, National Institute of Information and Communications Technology, Suita, as a Senior Researcher (Principal Investigator). He has also been affiliated as a Guest Professor with the Graduate School of Medicine and Graduate School of Frontier Biosciences, Osaka University. His research interest is the quantitative understanding of neural computation and representation.

Dr. Nishimoto's work was honored as one of Time Magazine's 50 Best Inventions of 2011, and he won the Ichimura Prize in Science for Distinguished Achievement in 2017.