

Claims

1. A method for generating a diagnostic question set of a data analysis framework for a new user, the method comprising:

step a of establishing, with a server, a question database including a plurality of questions, of collecting, with the server, solving result data of a plurality of users for the plurality of questions, and of applying, with the server, the solving result data to the data analysis framework, thereby calculating modeling vector(s) of the plurality questions and/or the plurality of users;

step b of extracting, with the server, from the question database, at least one candidate question from the plurality of questions for generating the diagnostic question set;

step c of identifying, with the server, an existing user from the plurality of users for whom solving result data for the at least one candidate question exists, and at least one other question for which solving result data of the existing user exists;

step d of applying, with the server, only the solving result data of the existing user for the at least one candidate question to the data analysis framework, thereby calculating a modeling vector of a virtual user;

step e of applying, with the server, the modeling vector of the virtual user, thereby calculating a virtual correct answer probability for the at least one other question;

step f of comparing, with the server, the virtual correct answer probability for the at least one other question with an actual solving result of the existing user for the at least one other question, and of

averaging, with the server, a comparison result according to the number of the plurality of existing users for whom solving result data for the at least one candidate question, and thereby calculating, with the
5 server a first average comparison result for the at least one candidate question; and

step g of, based on the first average comparison result for the at least one candidate question, generating, with the server, the diagnostic question set
10 for the new user.

2. The method as claimed in claim 1, further comprising:

generating, with the data analysis server, the
15 diagnostic question set comprising a plurality of additional candidate questions for which a corresponding average comparison result for each additional candidate question is within a threshold value of the first average comparison result.

20