A Survey on Cognitive Diagnosis Modeling

Jiaming Zhong

College of Information Science and Technology, Jinan University, Guangzhou 510632, China.

zhong_jia_ming@163.com

Abstract

Cognitive diagnosis is a fundamental problem in intelligent education, which aims to discover students' proficiency in specific knowledge concepts. Existing DINA cognitive diagnostic models have the advantages of high diagnostic accuracy and strong interpretability and have attracted the attention of a large number of relevant personnel. And most of the current CD-CAT systems use the DINA model as a diagnostic model. However, the DIAN model only performed simple 0-1 classification of the potential knowledge status of the subjects, which limited the application of the DINA model to multi-level scoring materials and also limited the use of CD-CAT in the field of education. Based on the DINA model, we propose a T-DINA model that is suitable for multi-level scoring data. And we implemented the parameter estimation program of the improving based the EM algorithm. The experimental results show that the T-DINA model has the characteristics of easy parameter estimation and high diagnostic accuracy.

Keywords

DINA, CD-CAT, EM algorithm.

1. Introduction

Cognitive diagnosis is a new method combining cognitive discipline and psychometric measurement[9]. It uses computers as the primary research tool to explore the cognitive processes and results of participants. The performance of the cognitive diagnostic model determines the predictive performance of CD-CAT. DINA model as a representative of cognitive diagnostic models has attracted the attention of many researchers due to its simple and precise mathematical definition, easy estimation of model parameters, and high interpretability. However, the traditional DINA model is only suitable for 0-1 second-level scoring data, which has poor compatibility with response data, which limits the practical application of the DINA model in the field of CD-CAT and cognitive diagnosis. The development of multi-level DINA models or other multi-level Cognitive diagnostic models has become an urgent need in the current field of cognitive diagnostics.

2. Literature Review

Cognitive diagnostic assessment(CDA)[3] was born in 1980.As of 2007, researchers have proposed more than 60 cognitive diagnostic models[5]. These cognitive diagnostic models had derived from the linear logistic trait model (LLTM)[4] proposed by Fisher and the rule space methodology (RSM)[11] proposed by Tatsuoka. Other cognitive diagnostic models, such as DINA (deterministic inputs, noisy "and" gate) model[7], HO-DINA (higher-order deterministic inputs, noisy "and" gate)[2] model, NIDA (noisy inputs, deterministic and gate) model, fusion model, GDM model (general diagnostic model)[6], GDM model (general diagnostic model)[12], etc. Besides, the Neural Cognitive Diagnosis Model (Neural CDM) based on deep neural networks has been born in the past two years[13].

Wang, Liu et al[13] proposed a general neurocognitive diagnostic framework. This framework automatically generates item response functions through deep learning and projects students and exercises into factor vectors. In the experiments, good results had achieved. Although the framework relies on the monotonic hypothesis of educational psychology, the item response function generated by the neural network has certain interpretability. However, compared with traditional cognitive diagnostic models, the interpretability of neurocognitive diagnostic models still has a large gap. Moreover, the comparative experiment of this model was still based on the idea of item response theory, and it compares merely the single scores of the participants, not the model accuracy rate. De la Torre[1] developed the parameter estimation procedures for the DINA and HO-DINA models based on the EM algorithm (expectation-maximization algorithm) and MCMC algorithm (Monte Carlo simulation and Markov Chain), and also gave the EM algorithm in DINA Derivation process in the model. At the same time, verification experiments prove that the EM algorithm and MCMC algorithm can be used as parameter estimation algorithms for cognitive diagnosis models.

3. DINA Model

3.1 Mathematical Definition

The mathematical expression of the DINA model is:

$$P\left(Y_{ij}=1\middle|\alpha_i\right)=\left(1-s_j\right)^{\eta_{ij}}{g_j}^{1-\eta_{ij}} \tag{1}$$

Equation 1 is also called the item response function of the DINA model. It describes the relationship between the probability that the subject i correctly answers the item j, the knowledge state α_i (attribute mode), and the item j parameters. Y_{ij} is the response data of the subject i to the item j, indicating whether the subject i responded correctly to the item j: $Y_{ij}=0$ indicates that the answer was incorrect, and $Y_{ij}=1$ indicates that the answer was correct. α_{i} represents the knowledge state of subject i. For items that examine K attributes, $\alpha_i=(a_1,a_2,\dots,a_{k-1},a_k)$， $a_i\in\{0{,}1\}$： $a_i=1$ represents the subject grasp of all attributes of item j, otherwise it is not grasped. $P\left(Y_{ij}=1\mid\alpha_i\right)$ is a posterior probability, which refers to the probability that the knowledge state of subject i is α_i, and subject i correctly answer the item j. η_{ij} is an indicator parameter, describing whether the subject i has all the knowledge (attributes) of item j's investigation:

$$\eta_{ij}=\prod_{k=1}^{K}{\alpha_{ik}}^{q_{jk}} \tag{2}$$

q_{jk} in equation 2 represents the relationship between item j and attribute k, and α_{ik} represents the relationship between subject i and attribute k. $\alpha_{ik}\in\{0{,}1\}$, $q_{jk}\in\{0{,}1\}$, α_{ik}=1 indicates that the subject i has grasped the attribute k. Otherwise it is not grasped; q_{jk}=1 indicates that the item j has inspected the attribute k. Otherwise it has not been inspected. $\eta_{ij}\in\{0{,}1\}$, which includes the relationship between subject i and item j. $\eta_{ij}=1$ indicates that subject i has grasped all the attributes of item j, and $\eta_{ij}=0$ indicates that subject i do not master all properties in item j.

The error parameter s_j for item j in equation 2 represents the probability that the subject answer the item j incorrectly under the premise of mastering all the attributes examined by the item j, which can be express as:

$$s_j=P\left(Y_{ij}=0\mid\eta_{ij}=1\right) \tag{3}$$

The guess parameter g_j for item j in equation 2 indicates that the subject does not fully grasp the attributes examined by item j, but the probability of correctly answering item j can be express as:

$$g_j=P\left(Y_{ij}=1\mid\eta_{ij}=0\right) \tag{4}$$

3.2 Advantages and Disadvantages

The mathematical definition of the DINA model is concise and clear. It assumes that each test attribute is independent, and there is no compensation effect. That is, to correctly answer the test item, the test subject must master all the attributes contained in the test item. Therefore, in the hypothesis of the DINA model, the potential knowledge status of the subjects was classified 0-1: mastering all the attributes investigated by the project, and at least one of the attributes investigated did not master.

That leading the subjects to a single test item only have two probabilities of correct answers: the probability of mastering all the attributes measured by the project without errors $(1-s_j)$, and the probability of not knowing all attributes of the projects but correct answering by guessing g_j.

The DINA model's assumptions about the knowledge state of the subjects are too simple. For a project examining the number of is K=5, the subjects who do not grasp one attribute and the subjects who have grasped the four attributes both answer correctly with the same guess probability g_j. In fact, in most cases, the more the subject grasp the attributes of the project, the higher the probability of correct responses. For example, there is an actual fractional subtraction item $3\frac{4}{5}-3\frac{2}{5}$. This item examines two attributes, attribute 1 is the basic fraction subtraction, and attribute 2 is the separation of integers from the score. Suppose there are two types of subjects answering this item. The first type of subjects have grasped attribute 1 but not attribute 2; the second type of subjects did not grasp both attributes. According to the DINA model, the two types of subjects answered correctly with the same probability. However, the probability of correct answer is higher in the first category than in the second category[8]. Studies by Sinharay[10], Fu[5], and others have shown that there is a significant difference in the probability of correct responses between participants who have grasped some attributes and those who do not have one attribute.

4. T-DINA Model

4.1 Mathematical Definition

The potential knowledge status of the subject are too simple in DINA model, especially the knowledge status of the subject is only final 0-1 classification. In fact, the number of subjects who had not grasped one attribute is tiny. Most of them are grasped at least one attribute of the project investigation. In view of this, this study proposes an improved T-DINA model based on the DINA model. This model assumes that subjects with different knowledge states have a total of three different probabilities of correctly answering the questions. These three probabilities correspond to the potential of the T-DINA model to the subjects. Three divisions of knowledge state: fully grasp the measured attributes of the project, partially grasp the measured attributes of the project, and entirely not grasp the measured attributes of the project. In the T-DINA model, fully grasping the measured attributes of the project means that the subject has grasped all the attributes measured by the project. Partially grasping the attributes of the project means that the subject has grasped at least one but not all of the attributes measured by the project. Completely not grasping the measured attributes of the project means that the subjects did not grasp any of the measured attributes of the project.

The mathematical expression of the T-DINA model is:

$$P(Y_{ij}=1 \mid \alpha_i) = (1-s_j)^{\eta_{ij}} {g_j}^{\gamma_{ij}} {c_j}^{\lambda_{ij}} \tag{5}$$

among them:

$$\eta_{ij} = \prod_{k=1}^{K} {\alpha_{ik}}^{q_{jk}} \tag{6}$$

$$\gamma_{ij} = \begin{cases} 1 & if\ \alpha_i q_j = 0 \\ 0 & if\ \alpha_i q_j \neq 0 \end{cases} \tag{7}$$

$$\lambda_{ij} = \begin{cases} 1 & if\ \alpha_i q_j \neq 0 \text{ and } \eta_{ij} = 0 \\ 0 & if\ \alpha_i q_j \neq 0 \text{ and } \eta_{ij} = 1 \end{cases} \tag{8}$$

$$s_j = P(Y_{ij}=0 \mid \eta_{ij}=1, \gamma_{ij}=0, \lambda_{ij}=0) \tag{9}$$

$$g_j = P(Y_{ij}=0 \mid \eta_{ij}=0, \gamma_{ij}=1, \lambda_{ij}=0) \tag{10}$$

$$c_j = P(Y_{ij}=0 \mid \eta_{ij}=0, \gamma_{ij}=0, \lambda_{ij}=1) \tag{11}$$

$P(Y_{ij}=1 \mid \alpha_i)$ in formula 5 refers to the probability that the subject i with knowledge state α_i correctly answers the item j. Let $\alpha_i=(\alpha_{i1},\dots,\alpha_{i(K-1)},\alpha_{iK})$, η_{ij} , γ_{ij} and λ_{ij} describe the potential knowledge status of subject i's investigation attributes relative to item j. $\alpha_i q_j$ is the inner product of the attribute vector of subject i and the attribute vector examined by item j. $\eta_{ij}=1$ means that subject i has grasped all the attributes of item j, and will correctly answer item j with a probability of $(1-s_j)$. $\gamma_{ij}=1$ means that subject i has not grasped all the attributes of item inspection and will use g_j probability answer item j correctly. $\lambda_{ij}=1$ means that the subject i has the attributes of the item inspection, and will answer the item j correctly with the probability of c_j. In the T-DINA model, c_j is a speculative parameter, which represents the probability that the subject will correctly answer the item on the premise that has partially grasped the attributes of the item investigation. η_{ij}, γ_{ij} and λ_{ij} can only have one equal one at the same time, and the others are 0. It is important to note that if only two attributes are examined in project j, then the T-DINA model will degenerate into the DINA model. In this case, the λ_{ij} parameter does not exist, and the T-DINA model and the DINA model are equal.

Under the assumption that the project response is locally independent, the likelihood function of the T-DINA model is consistent with the likelihood function of the DINA model as:

$$L(s,g;\alpha)=\prod_{i=1}^{N}\prod_{j=1}^{m}\left\{p_j(\alpha_i)^{Y_{ij}}\left[1-p_j(\alpha_i)\right]^{1-Y_{ij}}\right\} \tag{12}$$

In formula 12, N is the number of subjects, m is the number of attributes examined by item j. $p_j(\alpha_i)$ represents the probability that subject i will correctly answer item j on the premise that the knowledge state is α_i. Y_{ij} represents the result of the subject i answered the item j. $Y_{ij}=1$ indicates that the answer is correct. Otherwise, it is incorrect.

Under the assumption of the T-DINA model, it is necessary to estimate an extrapolated parameter c more than the DINA model for items with more than two attributes under consideration. Therefore, this model is more complicated than the traditional DINA model.

5. EM Algorithm Derivation Formula

Propose a cognitive diagnostic model, and its parameter estimation procedure must be implemented. Otherwise, it has only theoretical significance and cannot be applied to practice.Based on the EM algorithm, we implement the parameter estimation procedure of the T-DINA model according to the following formula:

$$s_j=\frac{I_{jl}^{(0)}-R_{jl}^{(0)}}{I_{jl}^{(0)}} \tag{13}$$

$$c_j=\frac{R_{jl}^{(1)}}{I_{jl}^{(1)}} \tag{14}$$

$$g_j=\frac{R_{jl}^{(1)}}{I_{jl}^{(1)}} \tag{15}$$

$I_{jl}^{(0)}$ represents the expectation of the number of subjects who have grasped all the test attributes of item j among all the participants. $R_{jl}^{(0)}$ represents the expectation of the number of participants who have grasped all the test attributes of item j and answered correctly. $I_{jl}^{(1)}$ represents the expectation of the number of participants who have grasped the properties measured by item j among all the participants. $R_{jl}^{(1)}$ represents the expectation of the number of participants who have grasped the measured attributes of item j and answered correctly in all the participants. $I_{jl}^{(2)}$ represents the expectation of the number of subjects who have grasped the measured attributes of 0 items j among all the subjects. $R_{jl}^{(2)}$ represents the expectation of the number of correctly answered attributes of 0 items j among all the participants.

6. Parameter Estimation Experiment

6.1 Experimental Data

Simulation conditions similar to de la Torre[1] were used.There are 30 test items and 5 test attributes. The Q matrix is shown in Table 1. The shape is 30 rows and 5 columns. The five attributes have 32 combined states. We simulate 30 subjects for each combined state, a total of 1920 subjects. Fix s = 0.2, c = 0.3, and g = 0.1 for each item. A total of 100 experiments were performed, and the average value was used as the final result.

Table 1. Q-Matrix for the Simulated Data

Item	Attribute						Attribute				
	1	2	3	4	5	Item	1	2	3	4	5
1	1	0	0	0	0	16	0	1	0	1	0
2	0	1	0	0	0	17	0	1	0	0	1
3	0	0	1	0	0	18	0	0	1	1	0
4	0	0	0	1	0	19	0	0	1	0	1
5	0	0	0	0	1	20	0	0	0	1	1
6	1	0	0	0	0	21	1	1	1	0	0
7	0	1	0	0	0	22	1	1	0	1	0
8	0	0	1	0	0	23	1	1	0	0	1
9	0	0	0	1	0	24	1	0	1	1	0
10	0	0	0	0	1	25	1	0	1	0	1
11	1	1	0	0	0	26	1	0	0	1	1
12	1	0	1	0	0	27	0	1	1	1	0
13	1	0	0	1	0	28	0	1	1	0	1
14	1	0	0	0	1	29	0	1	0	1	1
15	0	1	1	0	0	30	0	0	1	1	1

6.2 Metrics

ABSE reflects the degree of absolute deviation between the parameter estimate and the true value. The smaller the value, the more accurate the estimate, and it can examine the return to trueness or accuracy of the parameter estimate. ABSE mainly examines the trueness of the project parameters (s, c, g) in the experiment, and its calculation formula:

$$ABSE = \frac{\sum_{i=1}^{N} |y_i - \hat{y_i}|}{N}$$

N refers to the number of experiments. In this experiment, 100 experiments were simulated in 1920 subjects, N=100. y_i represents the true value of the parameter to be estimated, and $\hat{y_i}$ represents the estimated value of the parameter.

RMSE is the square root of the mean of the sum of the squared deviations of the true and estimated values in multiple experiments. It is particularly sensitive to a set of observed values that are too large or too small, and can well reflect the stability of the parameter estimation program. A smaller value indicates a more stable parameter estimation program. In this experiment, for 100 batches of different data to estimate the project parameters, RMSE can well reflect the stability of the EM estimation program. The calculation formula is as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$$

PMR(Pattern Match Ratio) represents the percentage of subjects who have correctly diagnosed the knowledge status of the subjects in the total number of subjects. It is a crucial evaluation indicator of the accuracy of the cognitive diagnosis model. The higher the value, the more accurate the model diagnosis is. In the field of cognitive diagnostic theory, PMR can examine the truthfulness of project estimation procedures. Calculated as follows:

$$MMR = \frac{N_p}{N}$$

N_p refers to the number of subjects correctly diagnosed by the model (each attribute of the subject was diagnosed correctly), and N represents the total number of subjects.

MMR(Marginal Match Ratio) refers to the accuracy rate of a single attribute:

$$MMR = \frac{1}{K}\sum_{k=1}^{K} MMR(k) = \frac{1}{K}\sum_{k=1}^{K}\frac{n_k}{N}$$

K is the number of test attributes, n_k represents the number of subjects with a correct diagnosis of the kth attribute, and N represents the total number of subjects. $MMR(k)$ represents the accuracy rate of the kth attribute, and MMR is the average of the accuracy rates of all attributes.

6.3 Experimental Results and Discussion

Table 2 is the estimation result of the parameters of the T-DINA model project, where "-" indicates that the item does not have corresponding parameters. For example, item 1 only examines one attribute and has only two states of complete mastery and no mastery at all. Therefore, the c parameter of this project is "-", which means that some of the potential knowledge states of the project's investigation attributes do not have this state. The results show that the average of the project parameters obtained by 100 repeated simulation estimates is very close to the true value. The average values of the project parameters of the 30 projects are 0.2003, 0.1002, and 0.2991, which are close to the project's true values of 0.2000, 0.1000, and 0.3000.

The ABSE of most project parameters is below 0.02, the ABSE of the inferred parameter c is significantly smaller than the ABSE of the incorrect parameter and the guessed parameter. The RMSE of the project parameters is similar to that of ABSE. The RMSE of the inferred parameter is significantly smaller than the RMSE of the incorrect parameter and the guessed parameter.

Table 2. T-DINA Model EM Algorithm Project Parameter Estimation Results

Item	Mean Estimate			Standard Error					
				ABSE			RMSE		
	s	g	c	s	g	c	s	g	c
1	0.2012	0.1010	-	0.0124	0.0092	-	0.0152	0.0119	-
2	0.2009	0.0992	-	0.0115	0.0097	-	0.0143	0.0119	-
3	0.1983	0.0980	-	0.0126	0.0089	-	0.0164	0.0115	-
4	0.1996	0.0970	-	0.0120	0.0102	-	0.0152	0.0128	-
5	0.2023	0.0989	-	0.0117	0.0099	-	0.0145	0.0120	-
6	0.2029	0.1016	-	0.0128	0.0083	-	0.0158	0.0099	-
7	0.2018	0.1015	-	0.0119	0.0090	-	0.0148	0.0111	-
8	0.2004	0.1012	-	0.0118	0.0098	-	0.0148	0.0124	-
9	0.2037	0.1013	-	0.0109	0.0096	-	0.0138	0.0120	-
10	0.2010	0.1011	-	0.0114	0.0097	-	0.0140	0.0122	-
11	0.2002	0.0985	0.3004	0.0161	0.0138	0.0118	0.0202	0.0175	0.0142

12	0.1981	0.0970	0.2977	0.0158	0.0139	0.0129	0.0199	0.0173	0.0164
13	0.2025	0.1001	0.2988	0.0183	0.0131	0.0150	0.0226	0.0163	0.0188
14	0.1979	0.1009	0.3007	0.0165	0.0134	0.0126	0.0209	0.0166	0.0157
15	0.1999	0.0994	0.2996	0.0181	0.0132	0.0122	0.0221	0.0170	0.0151
16	0.1985	0.0992	0.2972	0.0166	0.0135	0.0144	0.0210	0.0175	0.0181
17	0.1973	0.1013	0.2989	0.0178	0.0145	0.0149	0.0222	0.0178	0.0184
18	0.2005	0.1011	0.2965	0.0169	0.0142	0.0138	0.0209	0.0173	0.0173
19	0.2018	0.1004	0.2982	0.0183	0.0136	0.0128	0.0226	0.0167	0.0159
20	0.2048	0.0988	0.2988	0.0188	0.0134	0.0122	0.0221	0.0176	0.0155
21	0.1958	0.0999	0.3011	0.0274	0.0218	0.0106	0.0325	0.0258	0.0132
22	0.1965	0.1024	0.3015	0.0226	0.0198	0.0100	0.0281	0.0243	0.0123
23	0.1987	0.0974	0.2992	0.0206	0.0202	0.0113	0.0253	0.0240	0.0140
24	0.1978	0.1048	0.2994	0.0273	0.0210	0.0094	0.0350	0.0256	0.0117
25	0.1987	0.1037	0.2988	0.0225	0.0171	0.0111	0.0286	0.0210	0.0135
26	0.2018	0.1030	0.2976	0.0241	0.0198	0.0094	0.0297	0.0250	0.0120
27	0.2040	0.0971	0.2989	0.0255	0.0222	0.0109	0.0325	0.0278	0.0132
28	0.2014	0.1012	0.3009	0.0208	0.0189	0.0097	0.0262	0.0249	0.0118
29	0.2013	0.1013	0.2982	0.0228	0.0215	0.0101	0.0298	0.0266	0.0122
30	0.1995	0.0986	0.2993	0.0249	0.0207	0.0100	0.0314	0.0253	0.0133
M	0.2003	0.1002	0.2991	0.0177	0.0145	0.0118	0.0221	0.0180	0.0146
SD	0.0022	0.0019	0.0013	0.0051	0.0046	0.0018	0.0064	0.0056	0.0022

Table 3 shows the diagnosis results of the subjects' knowledge status. The boundary accuracy rate of each attribute of the subjects obtained by 100 repeated simulations reached more than 94%. The maximum value is 94.12%, the minimum value is 94.03%, the average value is 94.09%, and the average standard deviation is 0.0026. The average value of the pattern accuracy for 100 iterations is 75.73%, and the standard error is 0.0089.

Table 3. T-DINA Model Knowledge State Estimation Results

Repeat Test	MMR(k)					MMR	PMR
	α_1	α_2	α_3	α_4	α_5		
M	0.9403	0.9409	0.9414	0.9412	0.9409	0.9409	0.7573
SD	0.0057	0.0049	0.0051	0.0051	0.0051	0.0026	0.0089

Table 2 and Table 3 show that the parameter estimation program is accurate and stable for the project parameter estimation results of the T-DINA model, the project parameter estimation, and the subject knowledge state estimation are good.

7. Conclusion and Future Work

This paper proposes a multi-level scoring model T-DINA model based on the DINA model. At the same time, the parameter estimation procedure of the T-DINA model is given based on the EM algorithm. The final experiments show that the T-DINA model has the advantages of high diagnostic accuracy and easy parameter estimation.Our future research on cognitive diagnosis may proceed in the following directions:1) Cognitive diagnosis based on neural network. Adaptive learning of general item response functions through neural networks.2)Development and research of the CD-CAT system using the T-DINA model as a diagnostic model.3) Research on the performance comparison between the T-DINA model and other cognitive diagnostic models.

References

[1] Jimmy De La Torre. 2009. DINA model and parameter estimation: A didactic. Journal of educational and behavioral statistics 34, 1 (2009), 115–130.

[2] Jimmy De La Torre and Jeffrey A Douglas. 2004. Higher-order latent trait models for cognitive diagnosis. Psychometrika 69, 3 (2004), 333–353.

[3] Louis V. DiBello, Louis A. Roussos, and William Stout. 2006. A Review of Cognitively Diagnostic Assessment and a Summary of Psychometric Models. In Psychometrics, C.R. Rao and S. Sinharay (Eds.). Handbook of Statistics, Vol. 26.Elsevier, 979 – 1030. https://doi.org/10.1016/S0169-7161(06)26031-0

[4] Gerhard H Fischer. 1973. Te linear logistic test model as an instrument in educational research. Acta psychologica 37, 6 (1973), 359–374.

[5] Jianbin Fu and Yanmei Li. 2007. Cognitively diagnostic psychometric models: An integrative review. In annual meeting of the National Council on Measurement in Education, Chicago, IL.

[6] Sarah McConnell Hartz. 2002. A Bayesian framework for the unifed model for assessing cognitive abilities: Blending theory with practicality. Ph.D. Dissertation. ProQest Information & Learning.

[7] Brian W Junker and Klaas Sijtsma. 2001. Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. Applied Psychological Measurement 25, 3 (2001), 258–272.

[8] Robert J Mislevy, Russell G Almond, Duanli Yan, and Linda S Steinberg. [n. d.]. Bayes Nets in Educational Assessment:Where the Numbers Come From1. ([n. d.]).

[9] Paul D. Nichols, Susan F. Chipman, and Robert L. Brennan. 1995. Cognitively Diagnostic Assessment.

[10] Sandip Sinharay and Russell G Almond. 2007. Assessing ft of cognitive diagnostic models a case study. Educational and Psychological Measurement 67, 2 (2007), 239–257.

[11] Kikumi K Tatsuoka and Maurice M Tatsuoka. 1997. Computerized cognitive diagnostic adaptive testing: Effect onremedial instruction as empirical validation. Journal of Educational Measurement 34, 1 (1997), 3–20.

[12] M Von Davier. 2005. A general diagnosis model applied to language testing data. Educational Testing Service. Technical Report. Research Report.

[13] Fei Wang, Qi Liu, Enhong Chen, and Zhenya Huang. 2019. Interpretable Cognitive Diagnosis with Neural Network for Intelligent Educational Systems. arXiv preprint arXiv:1908.08733 (2019).