Online-Journals.org

iJET International Journal of Emerging Technologies in Learning

iJET | eISSN: 1863-0383 | Vol. 19 No. 7 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijet.v19i07.50943

PAPER

Hybrid Approach Using Multi-Relational Weighted Matrix Factorization (WMRMF) and Cohen's Kappa (*Sk*) to Refine Educational Items Clustering

Denon Arthur Richmond Gono[1](✉), Appoh Kouame[2], Konan Marcellin Brou[2], Bi Tra Goore[2]

[1]National Pedagogical Institute for Technical and Vocational Education, Abidjan, Côte d'Ivoire

[2]National Polytechnic Institute – Houphouet Boigny, Yamoussoukro, Côte d'Ivoire

arthur.gono@ipnetp.ci

ABSTRACT

In the context of adopting the competency-based approach (CBA) as a new teaching methodology in sub-Saharan countries, and particularly in Côte d'Ivoire, the development of learning content that is aligned with the economic, socio-cultural, and scientific needs of society is of paramount importance. Educational experts have therefore proposed competencies and associated tasks for educational programs. However, these learning contents often face issues of task redundancy. The present paper aims to address this problem by proposing a hybrid approach to educational item clustering, combining weighted multi-relational matrix factorization (WMRMF) and Cohen's Kappa (*Sk*) techniques (*Sk*-WMRMF). This approach takes into account not only student performance and achievements but also a novel reflexive relationship, "tasks-require–tasks". To evaluate the *Sk*-WMRMF approach, we conducted a survey among students in general secondary schools in Côte d'Ivoire. With an accepted task redundancy threshold of 0.185, an RMSE score of 0.198, and an improvement rate of 90.47% in the "tasks–skills" mapping, the results demonstrate that *Sk*-WMRMF enhances the elaboration of "task-skill" mappings. This not only improves learning content but also facilitates the updating of curricula in accordance with the CBA approach.

KEYWORDS

tasks–skills mapping, matrix factorization (MF), similarity measure, learning content, educational items clustering

1 INTRODUCTION

Assessing the effectiveness of educational systems has highlighted the necessity for reforms in educational programs [1]. In many French-speaking countries, these reforms have resulted in the adoption of the competency-based approach (CBA) [2]. Researchers such as Perrenoud [3] and Mansour [4] have dedicated their work to this pedagogical

Gono, D.A.R., Kouame, A., Brou, K.M., Goore, B.T. (2024). Hybrid Approach Using Multi-Relational Weighted Matrix Factorization (WMRMF) and Cohen's Kappa (*Sk*) to Refine Educational Items Clustering. *International Journal of Emerging Technologies in Learning (iJET)*, 19(7), pp. 92–103. https://doi.org/10.3991/ijet.v19i07.50943

Article submitted 2024-05-28. Revision uploaded 2024-07-16. Final acceptance 2024-07-17.

approach. Their studies have identified CBA as a catalyst for curriculum enhancement [5], addressing the issue of educational system efficiency. It initiates a paradigm shift in teaching methodologies and techniques, as well as in the content of learning development [6]. Nevertheless, a recent study conducted by the Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN (PASEC) reveals that Ivorian students exhibit deficiencies in the language and mathematical skills required at the primary school level [7]. This situation can be attributed, firstly, to the absence of a universal manual for assessing students. Each teacher proposes tasks for mastering various skills, which can be detrimental, as these tasks are often linked to the old pedagogical approach (objective-based teaching) and are now largely unsuited to mastering contemporary learning content. To address the issues raised by PASEC [7], authors such as Assielou et al., have suggested incorporating emotions into the development of learning strategies [8]. Similarly, authors such as Kouakou et al., [9] have studied microeconomic determinants, specifically the impact of school governance on student performance in sub-Saharan Africa. This work is related to our own, wherein we proposed tools to assist educational experts in developing learning content [10]. However, all these studies in response to PASEC fail to address the redundancy of tasks in learning content.

This redundancy could introduce bias in the clustering of educational items. Indeed, task redundancy in learning can have both positive and negative effects, depending on its application. Appropriate redundancy can enhance understanding by presenting information in various forms, which is beneficial for students with different learning styles. However, excessive redundancy can result in cognitive overload and inefficient learning. Therefore, the present paper aims not only to refine the clustering of educational items but also to determine the optimal threshold for redundant tasks to improve skill mastery.

2 STATE OF ART

The adoption of the CBA in the education system is intended to develop educational programs that meet the economic, socio-cultural, and scientific needs of society. To facilitate the transition to this pedagogical approach, tools for creating "task-skills" maps have been developed to assist educators and other stakeholders in the educational process.

In the literature, there are two main approaches to clustering educational items: those based on similarity measures and those based on models [11]. The assumption underlying similarity-based approaches is that students will perform similarly on tasks requiring similar skills, aiming to identify similarities between different pairs of tasks. In this context, Řihák and Pelánek [12] have identified six similarity measures specific to education. Their results, based on rand index calculation, clearly indicate that S*k* method provides superior results for clustering educational items. However, this method does not account for the concept of learning within the context of educational item clustering. Similarity measures in the educational field implicitly assume that the latent trait (specific skill mastery) remains constant during the learning activity that generated the performance. Consequently, students' performance on tasks associated with the same skill should be highly correlated. This assumption has motivated authors such as Nazaretsky et al., [13] to propose a new similarity measure called Kappa Learning, aimed at improving the Kappa method.

Model-based approaches aim to reduce the complexity of the problem by identifying latent factors underlying tasks. Numerous studies utilize techniques based on matrix factorization (MF) to construct task-skill mappings. In this context, Kim and Park [14] propose the sparse nonnegative matrix factorization (SNMF) method, which presents itself as a viable alternative for item clustering and yields better results compared to the K-Means method. Unlike the parsimony introduced by these authors in the NMF nonnegative matrix factorization (NMF) technique, Beheshti and Desmarais [15] introduce partial-order constraints in MF. Their approach is based on the assumption that if a difficult task is successfully completed, an easier one should be as well. Another MF technique, also based on the NMF technique, is proposed by Desmarais et al., [16]. This technique assumes that any missing skill will result in failure when performing a task. While NMF has proven effective for single-skill tasks, there is a need to consider multi-skill tasks when mapping tasks to skills. To address this, new MF techniques have been proposed to improve and overcome NMF's limitations, particularly the alternating least squares (ALS) factorization technique. The ALS technique addresses the issues of unique solutions and local minima. To further overcome these problems, Desmarais and Naceur [17] suggest using task-skill mappings proposed by experts. According to these authors, the Q-matrix predicted after minimization will diverge as little as possible from the initial matrix, making comparison with the expert's Q-matrix and its improvement more feasible.

In a recent study, we proposed the WMRMF technique as an alternative for clustering educational items [10]. Unlike the ALS technique, WMRMF relies on several relationships to predict experts' "task-skill" mappings. We obtained an RMSE error of 0.216, indicating a disparity between the Q-Matrix proposed by experts and the Q-Matrix predicted by the WMRMF technique. This disparity can be attributed to the redundancy of learning tasks in the mapping proposed by domain experts.

Unfortunately, none of the existing methods in the literature adequately address the issue of task redundancy when clustering educational items. Therefore, the present paper aims to resolve this problem by proposing a new approach to clustering educational items based on multi-relational matrix factorization and similarity measures.

3 RELATED WORKS

This paper centers on the utilization of the WMRMF technique and S*k* measure for the clustering of educational items. Specifically, this section elucidates the operational methodologies of WMRMF and S*k* within this context.

3.1 Weighted multi-relational matrix factorization approach

The WMRMF approach extends classic MF methods by incorporating various types of relationships or domain-specific information, such as student skills and grades, to enhance prediction accuracy across diverse contexts. This approach is described in [18]. In their study, the authors consider two primary relationships: "student performs task" and "student has acquired," aiming to predict the "task requires skill" relationship. The "student performs task" relationship is represented as a matrix $R \in \mathbb{R}^{S \times I}$ that records the scores obtained by students when performing tasks I. The mapping of "student–task" is grounded in skills identified in guides and programs, referred to as the performance matrix. The prediction of a student's performance on a task i is given by equation (1):

$$\hat{p}_{si} = \sum_{f=1}^{F} w_{1_{sf}} w_{2_{if}} = w_{1_s} w_{2_i}^T \tag{1}$$

The "student has skills" relationship is also represented as a matrix known as the learning matrix. In this matrix, denoted with K representing the skill set, various levels of skill mastery are defined. The prediction of the "task requires skill" relationship is given by equation (2):

$$o^{WMRMF} = \sum_{r=1}^{M} \Theta_r \sum_{(s,i)\in R_r} \left(R_{r_{si}} - w_{r1_s} w_{r2_i}^T \right)^2 + \lambda \left(\sum_{n=1}^{N} \| W_n \|_F^2 \right) \tag{2}$$

$$\text{with } R_r = \left\{ \left(E_{1r}; E_{2r} \right) \right\} (r = 1 \ldots M)$$

In this technique, we consider N types of entities linked $\{E_1, \ldots, E_N\}$ by M types of relationships, $\{R_1, \ldots, R_M\}$ which may be highly correlated with each other [19]. The matrices $W_1, W_2, \ldots, W_n$ $(n \in N)$ represent the model parameters. These parameters are generally learned by optimizing using the stochastic gradient descent technique [20]. A factor weight θ gives the importance of one relationship in relation to another. The prediction error e_{si} (see equation 3) is calculated as the difference between the actual performance value R_{si} and the predicted performance value $\hat{p}_{si}$ for each pair (s, i).

$$e_{si} = \left(R_{si} - w_{1_s} w_{2_i}^T \right) \tag{3}$$

The parameters of the WMRMF model are updated using equations (4) and (5) [21]:

$$w'_{r1_{Sk}} = w_{r1_{Sk}} + \beta \left(2\Theta_r e_{r_{si}} w_{r2_{ik}} - \lambda w_{r1_{Sk}} \right) \tag{4}$$

$$w'_{r2_{ik}} = w_{r2_{ik}} + \beta \left(2\Theta_r e_{r_{si}} w_{r1_{Sk}} - \lambda w_{r2_{ik}} \right) \tag{5}$$

In these relationships, factor β designates learning rate.

3.2 Cohen's Kappa measurement

Cohen's kappa is a dimensionless index that can be used to express agreement between two raters in a single number. Sometimes abbreviated Kappa *Sk*, it is a measure of inter-rater agreement for nominal scales [22]. The definition of kappa is given by equation (6) [23].

$$Kappa = \frac{P_o - P_e}{1 - P_e} \tag{6}$$

$$\text{With } P_o = \sum_{i=1}^{r} p_{ii} \text{ and } P_e = \sum_{i=1}^{r} p_i + P_{+i}$$

P_o represents the level of observed agreement and cases' proportion to which evaluators agree. And P_e represents the expected agreement level or expected agreement proportion by chance. Typically, in an educational item clustering context, tasks or items are defined as raters, students as subjects to be classified, and students' responses as classification results. Student responses to a task are formalized by correct or incorrect answers and scored 1 when the answer is correct and 0 when the answer is incorrect. Based on this definition, we consider a contingency table summarizing student responses to two different items (tasks): i and j. Suppose that n students have completed both items (tasks). Students' numbers in each cell are defined in Table 1.

Table 1. Contingency table for i and j items

		$Item_j$ Correct	$Item_j$ Incorrect	
$Item_i$	*Correct*	a	b	$a+b$
	Incorrect	c	d	$c+d$
		$a+c$	$b+d$	n

Notes: a – Students' number who correctly answered i and j. This is a point of agreement. b – Students' number who incorrectly answered i and correctly answered j. This is a point of disagreement. c – Students' number who correctly answered i and incorrectly answered j. This is a point of agreement. d – Students' number who incorrectly answered i and j, this is a point of agreement. $n = a + b + c + d$.

Cases' number on which evaluators agree equals $a + d$. This agreement is formalized by equation (7) [24].

$$P_o = \frac{a+d}{n} \tag{7}$$

The items are independent in the sense that each item independently 'evaluates' whether a student belongs to students' category of mastering a given skill. Thus, we could calculate the agreement's level that is expected by chance as a sum of the products of marginal probabilities.

$$P_e = \frac{(a+b)(a+c)+(b+d)(c+d)}{n^2} \tag{8}$$

Substituting P_o and P_e in equation (6) gives us equation (9).

$$T_{ij} = \frac{2(ad-bc)}{(a+b)(a+d)+(a+c)(c+d)} \tag{9}$$

4 PROPOSED APPROACH

We introduce a hybrid approach for clustering educational items, termed S*k*-WMRMF, combining WMRMF and S*k* techniques. This method builds upon previous studies outlined in [10] and [25]. To achieve our objectives, we utilized the same dataset employed in prior studies [10], calculated task similarities as assessed by domain experts in matrix (T), and conducted machine learning using the performance matrix (R), achievement matrix (A), and task similarity matrix (T) to predict the Q-Matrix.

4.1 Data set

To develop the dataset, we worked with eleven (11) mathematics teachers at Lycée Moderne Khalil in Daloa, Côte d'Ivoire, throughout the 2021–2022 school year. This collaborative effort spanned three months and involved active student participation in various learning activities. The participation facilitated the acquisition of a learning matrix, referred to as the acquisition matrix (A). Additionally, based on the curriculum guide and syllabus for lower 6th grade students developed by the ministry of education and literacy (MENA), these teachers devised tasks (tests, MCQs) designed to assess specified skills. These tasks (200) were administered to all students (200) during assessment sessions. The scores obtained by students across these tasks were used to construct the performance matrix (R), which comprises 40,000 performance entries.

As our main objective was to refine Q-Matrix's prediction, we also based out work on "task-skill" mapping proposed by experts in [10] by integrating task similarity, an important factor in learning's content development in CBA. All these matrices will be used as data sources for our proposed approach.

4.2 Determining task similarity

In addition to the relationships explored in [10], we consider a reflexive relationship between different tasks. This relationship defines similarity between tasks proposed by domain experts and is represented as a square matrix in which the rows and columns represent tasks performed by students. To calculate the similarity of two tasks q_1q_2, denoted $T_{q_1q_2}$ by the performance matrix, which summarizes different responses for these two tasks using the Kappa formula expressed in equation (9). Different student performances were obtained during various student assessments (tests, quizzes). Kappa is a value $T_{q_1q_2}$ for two tasks q_1 and q_2 is a number in a range [–1, 1]. Table 2 shows a set of Kappa values for seven tasks.

Table 2. Similarity matrix extract for seven tasks

Tasks		Tasks						
		q_{28}	q_{29}	q_{30}	q_{31}	q_{32}	q_{33}	q_{34}
	q_{28}	1	0.285	0.987	–0.363	0.285	0.545	0.285
	q_{29}	0.285	1	0.285	0.960	0.615	–0.666	–0.153
	q_{30}	0.987	0.285	1	0.285	–0.363	0.285	0.545
	q_{31}	–0.363	0.960	0.285	1	0.615	–0.666	–0.153
	q_{32}	0.285	0.615	–0.363	0.615	1	–0.92	–0.666
	q_{33}	0.545	–0.666	0.285	–0.666	–0.92	1	0.615
	q_{34}	0.285	–0.153	0.545	–0.153	–0.666	0.615	1

Similarity calculation results' analysis shows that tasks q_{28} and q_{30} are similar, with $T_{q_{28}q_{30}}$ coefficient's similarity value tending towards 1, while the coefficient's similarity for q_{29} and q_{33} tends towards –1.

4.3 Machine learning

For the learning phase, we used a performance matrix (R), an achievement matrix (A), and a task similarity matrix (T). The dataset consists of 40,000 (200 students × 200 tasks) performances, 16,800 (84 skills × 200 students) acquisitions, and 19,900 similarity measures. We used cross-validation, i.e., 4/5 for the learning phase and 1/5 for the testing phase.

The model parameters are:

- W_1 denotes a matrix in which each row s represents a vector containing F latent factors that best describe student's profile *s*.
- W_2 denotes a matrix where each row i is a vector containing F latent factors describing task *i*.
- W_3 denotes a matrix where each row c is a vector containing F latent factors describing skill *c*.

Proposed approach parameters (equations 4 and 5) were updated using algorithm below.

Algorithm 1: Stochastic Gradient Descent

1. Step 1: Parameter's initialization w_1, w_2 and w_3

2. Step 2: Calculation of prediction error

$e_{si} = \left(R_{si} - w_{1_s} w_{2_i}^T\right)$ (Equation 3)

3. Step 3: Update parameters

$w'_{r1_{Sk}} = w_{r1_{Sk}} + \beta\left(2\Theta_r e_{r_{si}} w_{r2_{ik}} - \lambda w_{r1_{Sk}}\right)$ (Equation 4)

$w'_{r2_{ik}} = w_{r2_{ik}} + \beta\left(2\Theta_r e_{r_{si}} w_{r1_{Sk}} - \lambda w_{r2_{ik}}\right)$ (Equation 5)

4. Step 4: Data assignment $(w_1, w_2) = (w'_{1_s}, w'_{2_i})$

5. Step 5: Repeat steps 2, 3 and 4 for another (*s*, *i*) pair.

In addition to the "student-performs-task" relationships, represented by a matrix $R \in \mathbb{R}^{S \times I}$ denoting the score obtained by students on task *I*, the "student-has-achievements" matrix shows mastery's level of the various K skills by the students, we also rely on the reflexive "tasks-requires-tasks" relationship, represented as a task matrix, $T \in \mathbb{R}^{T \times T}$ to predict the Q-Matrix $Q \in \mathbb{R}^{K \times I}$ (see Figure 1).

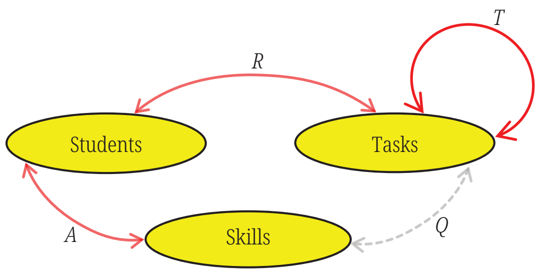

Fig. 1. Relationships taken into account in proposed approach

Concerning learning phase, we used a methodology based on the stochastic gradient descent algorithm through four steps:

- In step 1, we initialize model parameters w_1, w_2 and w_3 from normal distribution $N(\mu, \sigma^2)$ taking expectation $\mu = 0$ and standard deviation $\sigma = 0.01$.
- Next (step 2), we calculate prediction error for a given matrix (*R*, *A* and *T*) from equation (3).
- Then (step 3), we update model parameters from equations (4) and (5).
- Finally (step 4), for a pair (*i*, *j*) belonging to one of the matrices (*R*, *A*, and *T*) we repeat steps 2 and 3 until we have optimized the parameters w_1, w_2 and w_3.

To evaluate this approach, we used a 64-bit Windows environment, 16 GB RAM and an Intel Core i5 processor. Our algorithm was simulated using Python language.

5 RESULTS AND DISCUSSION

The study carried out in this paper aims to optimize educational item clustering. Unlike the work carried out in [10], we have extended the WMRMF technique to take into account the task's redundancy. To assess Q-Matrix's quality prediction, we used mean square error (RMSE). RMSE's expression is given by equation (10).

$$RMSE = \sqrt{\frac{\sum_{(r,s,i)\in D^{test}} \left(p_{si} - \hat{p}_{si}\right)^2}{\left|D^{test}\right|}} \tag{10}$$

With D^{test} corresponding to the *test* data set.

Thus, this section presents different RMSE results obtained from the WMRMF approach used in [10] for Q-Matrix prediction and from the "*Sk*-WMRMF" approach proposed in this paper.

5.1 Results

To evaluate the *Sk*-WMRMF approach, we relied primarily on *R, A* and *T* matrices. After evaluation, the parameters that optimize the model have been recorded in Table 3.

Table 3. Optimization parameters

Methods	Parameters
WMRMF	$K = 6$; $\#iter = 113$; $\beta = 10^{-2}$; $\lambda = 0.015$; $\Theta_R = 1$; $\Theta_A = 0.9$;
***Sk*-WMRMF**	$K = 5$; $\#iter = 149$; $\beta = 10^{-4}$; $\lambda = 0.015$; $\Theta_R = 1$; $\Theta_A = 0.9$; $\Theta_T = 1$

Figure 2 shows the RMSE results after 149 iterations. This error is calculated by equation (10) using the expert's Q-Matrix predicted by the WMRMF model and the Q-Matrix predicted by the *Sk*-WMRMF model.

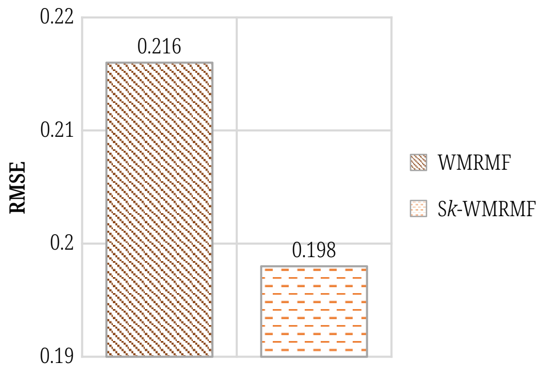

Fig. 2. Results for multi-relational weighted matrix factorization and cohen's kappa-multi-relational weighted matrix factorization models

We obtain an RMSE error = 0.216 for the WMRMF model from 113 iterations and an RMSE error = 0.198 for our proposed approach from 149 iterations. In Q-Matrix' prediction context, RMSE errors represent a form of disparity between basic Q-Matrix (Experts' Q-Matrix) and those predicted by the various models.

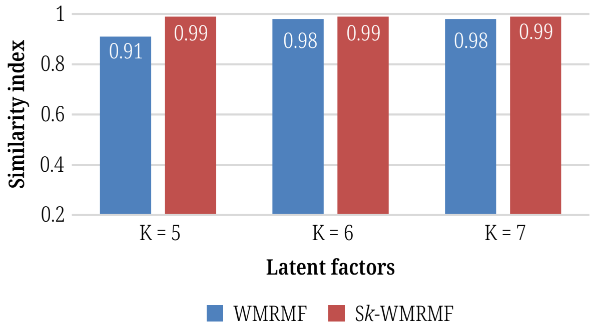

Fig. 3. Similarity indices between experts' Q-Matrix and predicted Q-Matrix by multi-relational weighted matrix factorization and cohen's kappa- multi-relational weighted matrix factorization models

For $K = 5$ and $K = 6$, the similarity index between the experts' Q-Matrix and the one predicted by our approach is 0.99, contrasting with 0.91 and 0.98 achieved using the WMRMF approach alone. Furthermore, the *Sk*-WMRMF approach identifies 168 irregularities out of 16,800 (84×200), compared to 378 irregularities found with the WMRMF approach alone. To validate the proposed *Sk*-WMRMF approach, we submitted the identified irregularities during evaluation to experts for their feedback. The analysis revealed that out of 168 proposals generated, experts deemed 152 acceptable, representing an improvement rate of 90.47% compared to 82.8% with the WMRMF technique alone. The results demonstrate that incorporating the reflexive relationship "tasks-requires–tasks" significantly enhances the prediction of the expert's Q-Matrix. It substantially reduces the number of irregularities submitted to experts for evaluation, effectively minimizing redundant tasks in the "tasks–skills" mappings.

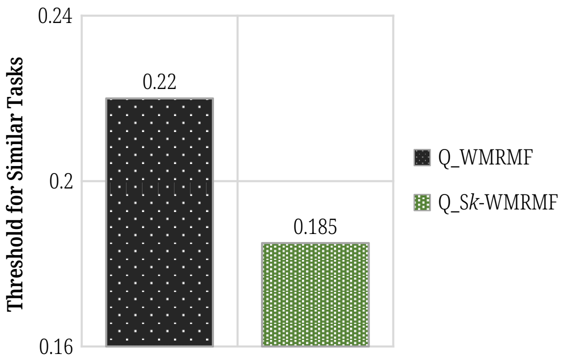

Fig. 4. Redundant task threshold allowed in expert Q-Matrices development for WMRMF and *Sk*-WMRMF models

Figure 4 illustrates the threshold of redundant tasks accepted in the development of the Q-Matrix. The rate of redundant tasks obtained from the Expert's Q-Matrix predicted by the WMRMF technique is 0.22, while it is 0.185 for the *Sk*-WMRMF approach. Given that the proposed approach has improved Q-Matrix prediction, the suggested threshold for better Q-Matrix prediction is 0.185.

5.2 Discussion

The results presented in Figures 2 and 3 clearly demonstrate that the S*k*-WMRMF model offers greater accuracy in elaborating task-skill mappings. This observation could be rationalized by the assumption that task redundancy in CBA learning content development can significantly influence the clustering of educational items. Moreover, it contributes to the development of learning strategies and skill mastery. According to Tagne et al., [26], the CBA aims to de compartmentalize content by linking strategies and related content to improve student participation and motivation. Unfortunately, social representations and practices associated with this pedagogical approach, coupled with the lack of learning manuals, still favor objective-based teaching as the preferred pedagogical model. This hypothesis is further supported by the results in Figure 4, which illustrate the threshold of redundant tasks accepted in the development of the experts' Q-Matrix. These results indicate that the number of redundant tasks introduces bias in the clustering of educational items.

6 CONCLUSION

In this paper, we propose a new approach to educational item clustering to address task redundancy in "task-skills" mappings. This approach is based on an ensemble technique that combines MF and S*k* measures. Unlike the WMRMF technique, which relies on the "student-performs-tasks" and "student-has-acquired" relationships, the S*k*-WMRMF technique adds the "tasks-requires–tasks" relationship. By applying the RMSE metric to our S*k*-WMRMF approach and conducting experiments on a dataset collected from a general secondary school in Côte d'Ivoire, we achieved an RMSE of 0.198. Additionally, out of 178 proposals generated by our approach, 152 were deemed acceptable, representing a 90.47% improvement rate in the experts' Q-Matrix. This result allows our approach to refine educational item clustering and consequently improve the learning content proposed within the CBA context. We also suggest a threshold of 18.5% redundant tasks for better Q-Matrix prediction. In future work, assessing the viability of the learning content offered by technical education, vocational training, and apprenticeship could help improve both the efficiency of the education system and the quality of the proposed workforce.

7 REFERENCES

[1] D. Lafontaine and M. Simon, "Évaluation des systèmes éducatifs," *Mes. évaluation en éducation*, vol. 31, no. 3, pp. 95–123, 2014. https://doi.org/10.7202/1024967ar
[2] I. Destin, "L'approche par compétences en contextes scolaires francophones: quels enjeux contextuels dans le cas d'Haïti et du Burkina Faso?" *Iramène Destin*, 2018.
[3] P. Perrenoud, "L'approche par compétences: une réponse à l'échec scolaire?" *Réussir au collégial*, vol. 71, p. 21356, 2000.
[4] A. Mansour, "Approche Par Compétences," *Reper.–IREM*, vol. 88, no. 1998, pp. 5–20, 2012.
[5] A. C. Pires-Ferreira, "Réforme du curriculum et approche par compétences au Cap Vert. Entre injonction internationale et projet national?" *Cah. la Rech. sur l'éducation les savoirs*, no. 13, pp. 183–210, 2014. https://doi.org/10.4000/cres.2651
[6] C. Chauvigné and J.-C. Coulet, "L'approche par compétences: un nouveau paradigme pour la pédagogie universitaire?" *Rev. française pédagogie*, no. 172, pp. 15–28, 2010. https://doi.org/10.4000/rfp.2169

[7] PASEC, *Résumé exécutif PASEC2019 Qualité du système éducatif en Afrique subsaharienne francophone : Compétences et environnement d'enseignement-apprentissage au primaire*, 2020.

[8] K. A. Assielou, D. A. R. Gono, K. D. Yao, and K. Sidibé, "Contribution de l'Intelligence Artificielle à la discrimination des émotions favorables à la performance des apprenants du secondaire second cycle de l'Enseignement Général," *Rev. RAMReS – Sci. Appl. l'Ing.*, vol. 5, pp. 48–56, 2023.

[9] A. K. Kouakou, S. Assienin, and K. C. N'da, "Do school governance factors determine student performance? A neural network analysis of data from French-speaking African countries," *SSRN Electron. J.*, 2023.

[10] D. A. R. Gono, B. T. Goore, Y. Tiecoura, and K. A. Assielou, "Multi-relational matrix factorization approach for educational items clustering," *International Journal of Information and Education Techmology (IJIET)*, vol. 13, no. 1, pp. 42–47, 2023. https://doi.org/10.18178/ijiet.2023.13.1.1778

[11] R. Pelanek, "Measuring similarity of educational items: An overview," *IEEE Trans. Learn. Technol.*, vol. 13, no. 2, pp. 354–366, 2019. https://doi.org/10.1109/TLT.2019.2896086

[12] J. Řihák and R. Pelánek, "Measuring similarity of educational items using data on learners' performance," in *Proc. 10th Int. Conf. Educ. Data Mining (EDM 2017)*, 2017, pp. 16–23.

[13] T. Nazaretsky, S. Hershkovitz, and G. Alexandron, "Kappa learning: A new item-similarity method for clustering educational items from response data," in *Proc. 12th Int. Conf. Educ. Data Min. (EDM 2019)*, 2019, pp. 129–138. [Online]. Available: https://files.eric.ed.gov/fulltext/ED599209.pdf

[14] J. Kim and H. Park, "Nonnegative matrix factorization for clustering," *Mod. Massive Data Sets*, pp. 1–26, 2008.

[15] B. Beheshti and M. Desmarais, "Improving matrix factorization techniques of student test data with partial order constraints," in *User Modeling, Adaptation, and Personalization, UMAP 2012,* in Lecture Notes in Computer Science, J. Masthoff, B. Mobasher, M. C. Desmarais, and R. Nkambou, Eds., Springer, Berlin, Heidelberg, vol. 7379, 2012, pp. 346–350. https://doi.org/10.1007/978-3-642-31454-4_33

[16] M. C. Desmarais, B. Beheshti, and R. Naceur, "Item to skills mapping: Deriving a conjunctive Q-matrix from data," in *Intelligent Tutoring Systems (ITS 2012),* in Lecture Notes in Computer Science, S. A. Cerri, W. J. Clancey, G. Papadourakis, and K. Panourgia, Eds., Springer, Berlin, Heidelberg, vol. 7315, 2012, pp. 454–463. https://doi.org/10.1007/978-3-642-30950-2_58

[17] M. C. Desmarais and R. Naceur, "A matrix factorization method for mapping items to skills and for enhancing expert-based Q-matrices," in *Artificial Intelligence in Education (AIED 2013),* in Lecture Notes in Computer Science, H. C. Lane, K. Yacef, J. Mostow, and P. Pavlik, Eds., Springer, Berlin, Heidelberg, vol. 7926, 2013, pp. 441–450. https://doi.org/10.1007/978-3-642-39112-5_45

[18] P. Nedungadi and T. K. Smruthy, "Personalized multi-relational matrix factorization model for predicting student performance," in *Intelligent Systems Technologies and Applications, Advances in Intelligent Systems and Computing*, S. Berretti, S. Thampi, and P. Srivastava, Eds., Springer, Cham, vol. 384, 2016, pp. 163–172. https://doi.org/10.1007/978-3-319-23036-8_15

[19] N. Thai-Nghe, M. Nhut-Tu, and H.-H. Nguyen, "An approach for multi-relational data context in recommender systems," in *Intelligent Information and Database Systems, ACIIDS 2017,* in Lecture Notes in Computer Science, N. Nguyen, S. Tojo, L. Nguyen, and B. Trawiński, Eds., Springer, Cham, vol. 10191, 2017, pp. 709–720. https://doi.org/10.1007/978-3-319-54472-4_66

[20] D. T. G. Takács, I. Pilászy, and B. Németh, "On the gravity recommendation system," *Processing*, vol. 2007, pp. 22–30, 2007.

[21] N. Thai-Nghe and L. Schmidt-Thieme, "Multi-relational factorization models for student modeling in intelligent tutoring systems," in *2015 IEEE Seventh Int. Conf. Knowl. Syst. Eng. (KSE)*, 2015, pp. 61–66. https://doi.org/10.1109/KSE.2015.9

[22] G. Rau and Y.-S. Shih, "Evaluation of Cohen's kappa and other measures of inter-rater agreement for genre analysis and other nominal data," *Journal of English for Academic Purposes*, vol. 53, p. 101026, 2021. https://10.1016/j.jeap.2021.101026

[23] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and Psychological Measurement*, vol. 20, no. 1, pp. 37–46, 1960. https://10.1016/10.1177/001316446002000104

[24] M. J. Warrens, "Five ways to look at Cohen's Kappa," *J. Psychol. Psychother.*, vol. 5, no. 4, pp. 1–4, 2015. https://doi.org/10.4172/2161-0487.1000197

[25] J. Sim and C. C. Wright, "The Kappa statistic in reliability studies: Use, interpretation, and sample size requirements," *Phys. Ther.*, vol. 85, no. 3, pp. 257–268, 2005. https://doi.org/10.1093/ptj/85.3.257

[26] G. Tagne, D. Asoh, and C. Gauthier, "Réussite scolaire et réforme éducative selon l'Approche Par Compétences (APC) en Afrique Subsaharienne : résultats d'une étude comparative," *Rev. l'Association Francoph. Int. Rech. Sci. en Educ.*, vol. 16, pp. 76–97, 2016.

8 AUTHORS

Denon Arthur Richmond Gono, PhD in Computer Science is a teacher-researcher at the National Institute of Technical and Vocational Education, Abidjan, Côte d'Ivoire. His areas of research include learning content design, performance prediction and learning assessment. He is a member of the Data Science and AI Laboratory (LASDIA) (E-mail: arthur.gono@ipnetp.ci).

Appoh Kouame, PhD in Computer Science is a lecturer and researcher at the Institut National Polytechnique Houphouet-Boigny, Yamoussoukro, Côte d'Ivoire. His areas of research include ontology, expert systems, traditional medicine and Artificial Intelligence. He is a member of the Computer Networks and Telecommunications Laboratory (LARIT).

Konan Marcellin Brou is Senior Lecturer at the Institut National Polytechnique Houphouet – Boigny. His research interests include distributed computing systems, expert systems, software engineering and artificial intelligence. He is a member of the Data Science and AI Laboratory (LASDIA) and head of the Digital & Industry team (Datalab).

Bi Tra Goore is Full Professor at the National Polytechnic Institute Houphouet – Boigny. His areas of research are the design, modeling and performance evaluation of sensor networks. He is Director of the Data Science and AI Laboratory (LASDIA).