Proceedings of the 3rd International Conference on Mathematical Physics and Computational Simulation
DOI: 10.54254/2753-8818/100/2025.21957

Application of Change Point Detection in the Rasch Model

Hongjie He

Department of Physics, Mathematics, and Computing,The University of Western Australia, Perth, Australia
24462401@student.uwa.edu.au

Abstract: This study introduces a novel approach to enhancing the traditional Rasch model by integrating Maximum Likelihood Estimation (MLE), Change Point Detection (CPD), and KL divergence. The modified Rasch model is designed to dynamically evaluate student performance, overcoming limitations of static parameter assessments by identifying shifts in students' abilities over time. MLE is used to accurately assess student responses, while CPD identifies significant changes in performance, potentially indicating shifts such as learning improvements or dishonest behavior. KL divergence further analyzes these shifts by measuring the variability in accuracy distributions between consecutive assessments. Applied to a synthetic dataset, this method demonstrates its potential in offering a more dynamic and detailed understanding of educational assessments. However, challenges such as data sparsity potentially affect the stability and accuracy of the model. Despite these limitations, this approach provides a promising framework for future enhancements in educational evaluation systems, indicating significant shifts in student abilities that traditional methods may overlook.

Keywords: Rasch Model, Maximum Likelihood Estimation, Change-Point Detection, Educational Assessment

1. Introduction

The application of the Rasch model in educational assessment has been a standard approach to measure students' abilities and the difficulties of questions effectively. Originally developed by Georg Rasch, this probabilistic model has been instrumental in the field of educational assessment, providing insights into the interactions between test takers and test items [1]. However, the traditional Rasch model primarily quantifies static parameters, representing them through the accuracy of responses to questions, which does not entirely capture the dynamic nature of a student's abilities that may fluctuate due to various factors such as learning progress, test-taking strategies, or external influences, including potential dishonest behaviors [2].

Recognizing the limitations of the static approach, there is a pressing need for a methodology that can adapt and respond to the changing dynamics of student performance over time [3]. Most existing change-point detection (CPD) techniques focus on time-series analysis, often overlooking the structured interactions inherent between students and questions, which are critical in educational settings [4-6]. Recent advancements suggest that graph-based methodologies could offer a superior representation of shifts in student abilities, providing a more nuanced approach to detecting changes [7-8].

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/100/2025.21957

This study proposes an innovative integration of Maximum Likelihood Estimation (MLE), Change Point Detection (CPD), and Kullback-Leibler (KL) divergence within the framework of the Rasch model to address these challenges. By enhancing the traditional model, this approach aims to dynamically assess shifts in student abilities and the nuances of question difficulty over time. MLE is employed to refine the accuracy of student response assessments, while CPD, aided by KL divergence, is utilized to identify and analyze significant changes in performance metrics, which may indicate learning advancements or irregular behaviors such as cheating [9]. This method not only enhances the detection of changes in educational assessments but also incorporates a more detailed analysis through the use of synthetic data sets, thereby testing the effectiveness and robustness of the proposed modifications [10].

This enhanced approach seeks to contribute to the educational field by providing tools that can dynamically adapt to the complexities of measuring educational achievement, thus offering a more accurate, timely, and comprehensive analysis of student performance. The application of this novel methodology is expected to yield insights that could lead to significant improvements in educational assessment practices, ultimately influencing teaching strategies and educational policy development.

2. Theoretical Background

2.1. The Rasch Model

The Rasch model is a probabilistic latent trait measurement model proposed by Geog Rasch, which mainly used to evaluate the relationship between individual abilities and the difficulty of test items [1].

$$\mathrm{P}(X_{ij}=1)=\frac{e^{(\theta_i-\beta_j)}}{1+e^{(\theta_i-\beta_j)}} \tag{1}$$

To address the limitations of static assumptions, there are many extensions of the Rasch model have been proposed. For example, the random pairing MLE (RP-MLE) and sparse Rasch models improve the estimation process by using the calculations of rolling accuracy when the response data is complete. This study also calculate the rolling accuracy and have a extra improvement, which is remove the 0/1 level. In fact, students often do some parts of one question correctly, not all of them wrong (0) or all of them right (1) in the real world. Therefore, the method changes the traditional Rasch model to a model with five accuracy levels.

$$\mathrm{P}(X_{ij}=x)=\mathrm{f}(\theta_i-\beta_j,x)(x\in\{0,0.25,0.5,0.75,1\}) \tag{2}$$

The quantile mapping method was used in this experiment to determine the more precise probability of each students' accuracy.

2.2. Maximum Likelihood Estimation (MLE) for Ability Estimation

MLE (maximum likelihood estimation) is a core method used in statistics for parameter estimation, which can find the parameter values that are most likely to generate observational data [2].

$$\mathrm{L}(\theta,\beta)=\sum_{i=1}^{N}\sum_{j=1}^{M}\sum X_{ij}\log\mathrm{P}\big(X_{ij}=1\big)+(1-X_{ij})\log\left(1-P(X_{ij}=1)\right) \tag{3}$$

MLE has been used in latent trait modeling with in item response theory (IRT) [6]. Recently, some research use the traditional MLE in the graph-based student model to highlight the tracks of potential performance of the students. However, in this study the MLE model has been changed following the Rasch model. The changed MLE has five levels of accuracy and can provide better ability to estimate the students' ability.

DOI: 10.54254/2753-8818/100/2025.21957

$$L(\theta, \beta) = \sum_{i=1}^{N} \sum_{j=1, x\epsilon\{0,0.25,0.75,1\}}^{M} \sum X_{ij} \log P(X_{ij} = x) \quad (4)$$

Additionally, the MLE estimation values in the experiment were calculated over multiple time intervals, allowing for a more detailed analysis of the changes in capability over time [3].

2.3. Change-Point Detection (CPD) Using KL Divergence

Kullback-Leibler Divergence (KL divergence) is an indicator that can measure the difference between two probability distributions and has non negativity and asymmetry.

$$D_{KL}(P||Q) = \int_{-\infty}^{\infty} p(x) log \frac{p(x)}{q(x)} dx \quad (5)$$

The KL divergence has been widely applied in network-based CPD tasks to capture the students' response distribution. Each students' response history is divided according to time intervals, then the distribution's analysis has been done in the consecutive intervals. If the score of KL divergence is very high, the students' response distribution must change, which may be influenced by many external or internal factors like severe and sustained illness or potential cheating behavior. This experiment sets a KL divergence threshold to calculate the distribution that students' accuracy for answering the questions. By doing so, the mutation points of students' abilities can be determined.

3. Experimental Methodology

3.1. Data Generation and Preprocessing

First and foremost, the experiment generate the random synthetic data, simulating 50 students answer 50 questions. The accuracy of response has been divided to 5 levels (0, 0.25,0.5,0.75,1) , which can reflect the students' abilities in different levels. Then, the code generate timestamps to ensure the temporal sequence of the detection data. To extend the analysis beyond individual students' accuracy, a bipartite graph has been created to simulate that 10 students answer 10 questions (These students and questions are randomly chosen by the primary synthetic data), which can represent the interactions between students and questions.

3.2. Experiment process (code process)

Firstly, the code generate the random data and timestamps with 5 levels' accuracy. After sorting by time, the code calculate the basic parameter (θ and β). Then, calculating the probabilities of the five accuracy levels and maping them to the original probabilities (0-1) of the basic Rasch model. Subsequently, the code randomly allocate the five accuracy levels based on probability's allocation. Then, using the MLE to calculate the special parameter of Rasch model, which should consider the 5 levels' accuracy.

Secondly, the code initialize parameters and add Rasch estimation values, which prepare for calculating rolling accuracy values. Afterwards, the code calculate the KL divergence, which based on the estimated ability values (accuracy values). Then, the code establish a KL divergence threshold (KL divergence more than 0.3). Additionally, the code draw the rolling ability change curve and mark the change points.

Thirdly, the code using the chosen data (10 students answer 10 questions) to draw the bipartite graph, which based on the accuracy calculated by the previous section.

Finally, the code draw the ROC curve for the whole experiment process and calculate the AUC - ROC score. And the code in this experiment can be found in the git-hub (https://github.com/Icy-north-ocean/Application-of-change-point-detection-in-the-Rasch-model).

DOI: 10.54254/2753-8818/100/2025.21957

4. Experimental Result

4.1. Change-Point Detection Results

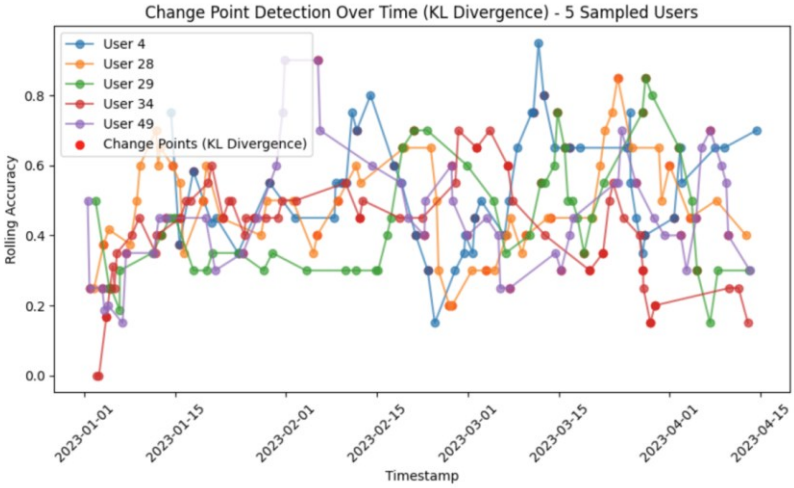


Figure 1: The graph of CPD results. (Picture credit : Original)

Figure 1 is the graph of CPD results, the study choose 5 students' data to draw the graph for the definition. The horizontal axis is Timestamps (from 2023-01-01 to 2023-04-15), and the vertical axis is Rolling Accuracy. There are many change points that one color corresponds to one student in the graph. By using the trend of rolling accuracy, the changes of each student's ability can be known and some abnormally high or low values may reveal the potential cheating behaviors or poor learning status of students.

4.2. Bipartite Graph Results

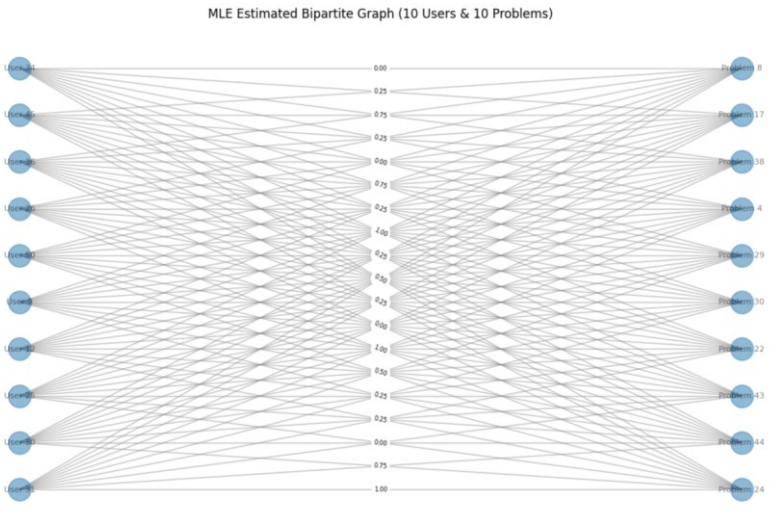


Figure 2: The MLE estimated bipartite graph. (Picture credit : Original)

Figure 2 is the MLE estimated bipartite graph. In the graph, 10 students in the left and 10 questions in the right. Each student has an independent accuracy (0, 0.25, 0.5, 0.75, 1) rate for each questions. This mechanism of answering extends the traditional Rasch model's function.

DOI: 10.54254/2753-8818/100/2025.21957

4.3. Performance Metrics

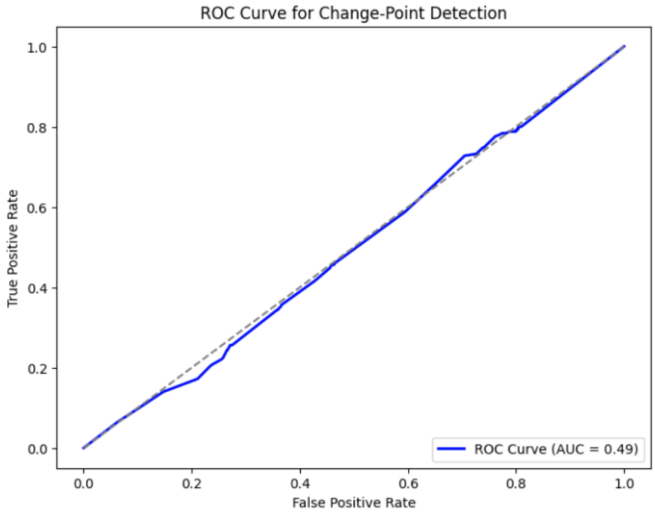


Figure 3: ROC curve and AUC score for CPD. (Picture credit : Original)

Figure 3 is the ROC curve of CPD. To evaluate the effectiveness of KL divergence in change-point detection, the AUC - ROC score was computed. The obtained AUC - ROC score is 0.4938.

5. Challenges

Through the experimental results, there are some problems can be found. Firstly, the amount of the random is small, so it may lead to over-fitting and it can't greatly stimulate the students and questions in the real world. Secondly, the AUC score is only 0.4938, which reveals that the model's accuracy is low.

In order to solve these shortcomings, future experiments and research on the first issue need to use large datasets like some real data records of students' accuracy in one course in the different time. For the second problem, Bayesian inference can be used to reduce the over-fitting through regularization. In addition, the JS divergence can substitute the KL divergence to improve the accuracy of the model and better detect the sudden changes in students' abilities.

6. Conclusion

This study aims to address the two problems in traditional Rasch model that it only have static parameter and simple form of accuracy (0, 1). Therefore, this study applies change point detection (CPD) to the Rasch model, calculates the parameters using maximum likelihood estimation, and sets timestamps for time sorting. Then, modifications are made to the traditional Rasch model and basic MLE formula by using probability reflecting to change the basic accuracy to a new five level accuracy. Afterwards, the experiment has been done on the computer with python code, which is designed for the study. Eventually, the output have three figures, which is change point detection result graph, bipartite graph and the AUC - ROC graph. The result graph shows the distribution of different students' accuracy over time, achieving the experimental goal. Overall, this study explores new directions for the application of change point detection and optimizes the traditional Rasch model.

References

[1] *Sulem, D., Kenlay, H., Cucuringu, M., & Dong, X. (2024). Graph similarity learning for change-point detection in dynamic networks. Machine Learning, 113, 1–44.*

DOI: 10.54254/2753-8818/100/2025.21957

[2] Yang, Y., & Ma, C. (2024). Random pairing MLE for estimation of item parameters in Rasch model. Department of Statistics, University of Chicago. arXiv preprint arXiv:2406.13989.

[3] Zhou, Z.-H., & Liu, F. (2020). Isolation Forest. In Proceedings of the 8th IEEE International Conference on Data Mining (pp. 413–422). IEEE.

[4] Zhu, X., Liu, Z., Cambria, E., Yu, X., Fan, X., Chen, H., & Wang, R. (2025). A client–server based recognition system: Non-contact single/multiple emotional and behavioral state assessment methods. Computer Methods and Programs in Biomedicine, 260, 108564.

[5] Wang, R., Zhu, J., Wang, S., Wang, T., Huang, J., & Zhu, X. (2024). Multi-modal emotion recognition using tensor decomposition fusion and self-supervised multi-tasking. International Journal of Multimedia Information Retrieval, 13(4), 39.

[6] Wang, F., Ju, M., Zhu, X., Zhu, Q., Wang, H., Qian, C., & Wang, R. (2025). A Geometric algebra-enhanced network for skin lesion detection with diagnostic prior. The Journal of Supercomputing, 81(1), 1-24.

[7] Zhao, Z., Zhu, X., Wei, X., Wang, X., & Zuo, J. (2021, June). Application of Workflow Technology in the Integrated Management Platform of Smart Park. In 2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC) (Vol. 4, pp. 1433-1437). IEEE.

[8] Zhang, Y., Zhao, H., Zhu, X., Zhao, Z., & Zuo, J. (2019, October). Strain Measurement Quantization Technology based on DAS System. In 2019 IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC) (pp. 214-218). IEEE.

[9] Bradley, R. A., & Terry, M. E. (1952). Rank analysis of incomplete block designs. Biometrika, 39(3/4), 324–345.

[10] Wang, X., & Li, J. (2024). Wasserstein distance for robust change-point detection. Neural Computation, 36(1), 178–198.