

INTISARI

Dalam dunia pendidikan, salah satu metode untuk mengukur tingkat kemampuan peserta didik adalah dengan melakukan ujian. Seiring perkembangan teknologi, metode pelaksanaan ujian terus diteliti dan mengalami inovasi. Salah satunya adalah pengembangan *computerized adaptive testing* yang memberikan pertanyaan secara satu persatu berdasarkan jawaban sebelumnya. *Graph-enhanced Multi Objective for Computerized Adaptive Testing* (GMOCAT) adalah salah satu penelitian di bidang CAT. GMOCAT menggunakan *reinforcement learning* untuk memilih pertanyaan berdasarkan tiga tujuan, yaitu kualitas, keberagaman, dan kebaruan. GMOCAT memiliki performa unggul dibandingkan penelitian-penelitian sebelumnya dan telah diuji pada dataset Eedi, Juni, dan ASSISTments2009. Akan tetapi, pemberian aturan keberagaman yang hanya berdasar pada kebaruan konsep yang dibawa, dapat mengakibatkan model cenderung memberikan konsep baru terus tanpa mempertimbangkan apakah model sudah yakin dengan estimasi kemampuan peserta didik. Penelitian ini bertujuan untuk melakukan modifikasi pemberian aturan reward keberagaman sekaligus modifikasi *termination criterion*. Selain itu, ditambahkan antarmuka simulasi ujian yang dapat memberikan gambaran pelaksanaan ujian. Model hasil modifikasi diuji menggunakan dataset DBEKT22 dan dianalisis performanya dengan metrik AUC, ACC, dan *Coverage*. Hasil penelitian ini meliputi: (1) terjadi penurunan sebesar 0,5% pada metrik ACC akibat modifikasi *termination criterion* dan *reward* keberagaman, (2) tidak terjadi perubahan signifikan pada metrik AUC akibat modifikasi, (3) *coverage* hanya mencapai 51,3% sebelum modifikasi dan 42,1% setelah modifikasi jika ditinjau pada langkah ke-20 akibat kurangnya data respons tiap peserta didik, dan (4) hasil simulasi salah satu peserta menunjukkan *coverage* dapat mencapai 97,46%.

Kata kunci: *computerized adaptive testing, neural cognitive diagnosis, reinforcement learning.*

ABSTRACT

Assessment is a crucial method for evaluating students' abilities in education. With technological advancements, examination methods have evolved, leading to innovations such as computerized adaptive testing (CAT), which presents questions sequentially based on previous responses. One notable study in this domain is the Graph-Enhanced Multi-Objective for Computerized Adaptive Testing (GMOCAT), which leverages reinforcement learning to select questions based on three objectives: quality, diversity, and novelty. GMOCAT has demonstrated superior performance compared to prior studies and has been tested on datasets like Eedi, Junyi, and ASSISTments2009. However, its diversity criterion, which prioritizes introducing novel concepts without considering the confidence level in estimating students' abilities, presents a limitation. This study aims to address this by modifying the diversity reward mechanism and the termination criterion. Additionally, a simulation interface for exams is developed to illustrate the testing process. The modified model was evaluated on the DBEKT22 dataset, with performance analyzed using AUC, ACC, and Coverage metrics. Results indicate: (1) 0.5% reduction in ACC, (2) no significant changes in AUC metrics following the adjustments, (3) the coverage only reached 51.3% before modification and 42.1% after modification when observed at step 20, due to the lack of response data for each participant, and (4) the simulation results for one participant indicate that the coverage can reach up to 97.46%.

Keywords : *computerized adaptive testing, neural cognitive diagnosis, reinforcement learning.*