



INTISARI

Perkembangan teknologi informasi di era digital membawa transformasi besar dalam bidang pendidikan. Kemudahan akses informasi dan adanya pandemi COVID-10 menambah alasan terjadinya pembelajaran secara daring. Peningkatan pengguna sistem pembelajaran daring mendorong inovasi untuk membuat sistem pembelajaran yang lebih terpersonalisasi. Salah satu inovasinya adalah *Intelligent Tutoring System (ITS)* yang berbasis *knowledge tracing* yang memungkinkan sistem untuk memodelkan status kognitif siswa dan melacak perubahan pengetahuan mereka seiring berjalannya waktu.

Pendekatan probabilistik masih menjadi pilihan utama dalam studi *knowledge tracing*, tapi akurasi masih terbatas. DKT sebagai pendekatan *deep learning* memiliki keunggulan dalam segi performa dan akurasi model yang lebih baik. Pada penelitian sebelumnya, AUC yang dihasilkan model DKT pada dataset ASSISTments 2019-2020 ada pada angka 80% yang mengindikasikan adanya ruang untuk optimalisasi. Kinerja DKT yang belum maksimal tersebut salah satu faktornya adalah kurangnya proses *pre-processing* yang mendalam, mengingat dataset ini memiliki percampuran antara data *skill builder* dan *non skill builder* yang berpotensi membuat *noise* dan mengurangi akurasi model. Selain itu, model DKT juga memiliki sensitivitas pada konfigurasi *hyperparameter* yang digunakan, penggunaan satu metode *hyperparameter tuning* seringkali tidak cukup untuk menentukan konfigurasi *hyperparameter* terbaik yang menghasilkan nilai AUC paling optimal.

Penelitian ini bertujuan untuk meningkatkan akurasi model DKT melalui proses *pre-processing* yang lebih mendalam—termasuk pemisahan tipe data dan modifikasi nilai *skill*—serta analisis komparatif metode *hyperparameter tuning* menggunakan Random Search, Bayesian Optimization, dan Hyperband. Hasil penelitian menunjukkan bahwa konfigurasi *hyperparameter* terbaik dihasilkan oleh Bayesian Optimization (*batch size* 20, LSTM *units* 250, *dropout rate* 0.3, dan *optimizer* Adam) dengan capaian nilai AUC sebesar 97,95%. Capaian ini merepresentasikan kenaikan performa sebesar 17% dibandingkan *baseline* penelitian sebelumnya. Dengan demikian, penelitian ini memberikan kontribusi dalam membuktikan bahwa kombinasi pembersihan data pada dataset ASSISTments 2019-2020 serta pemilihan metode optimasi yang tepat sangat penting untuk memaksimalkan performa model DKT dalam memprediksi pengetahuan siswa.

Kata kunci : *deep learning, knowledge tracing, deep knowledge tracing, hyperparameter tuning*



ABSTRACT

The development of information technology in the digital era has brought major transformations in the field of education. The ease of access to information and the COVID-19 pandemic have further reinforced the reasons for the adoption of online learning. The increase in users of online learning systems encourages innovations to create more personalized learning systems. One such innovation is the Intelligent Tutoring System (ITS) based on knowledge tracing, which allows the system to model students' cognitive status and track changes in their knowledge over time.

Probabilistic approaches remain the primary choice in knowledge tracing studies, yet their accuracy is still limited. Deep Knowledge Tracing (DKT), as a deep learning approach, offers advantages in terms of better performance and model accuracy. In previous research, the Area Under the Curve (AUC) generated by the DKT model on the ASSISTments 2019-2020 dataset was around 80%, indicating room for optimization. One factor contributing to this suboptimal DKT performance is the lack of in-depth pre-processing, considering that this dataset contains a mixture of “skill builder” and “non-skill builder” data, which has the potential to create noise and reduce model accuracy. Furthermore, the DKT model is sensitive to the hyperparameter configuration used; the use of a single hyperparameter tuning method is often insufficient to determine the best hyperparameter configuration that yields the most optimal AUC value.

This research aims to improve DKT model accuracy through a more in-depth pre-processing process—including the separation of data types and modification of skill values—as well as a comparative analysis of hyperparameter tuning methods using Random Search, Bayesian Optimization, and Hyperband. The results indicate that the best hyperparameter configuration was produced by Bayesian Optimization (batch size 20, 250 LSTM units, dropout rate 0.3, and Adam optimizer) achieving an AUC value of 97.95%. This achievement represents a performance increase of 17% compared to the baseline of previous research. Thus, this study contributes by demonstrating that the combination of specific data cleaning on the ASSISTments 2019-2020 dataset and the selection of appropriate optimization methods is crucial for maximizing DKT model performance in predicting student knowledge.

Keywords : deep learning, knowledge tracing, deep knowledge tracing, hyperparameter tuning