

ABSTRAK

Energi surya bersifat intermiten sehingga memerlukan mekanisme pengendalian orientasi panel yang adaptif agar perolehan energi tetap optimal. Penelitian ini mengembangkan sistem optimalisasi orientasi panel surya dengan mengintegrasikan *Temporal Fusion Transformer* (TFT) sebagai model peramalan daya dan *Double Dueling Deep Q-Network* (D3QN) sebagai agen pembelajaran penguatan. Data primer dikumpulkan dari 13 orientasi panel tetap dan sembilan variabel lingkungan selama empat minggu, kemudian direduksi dimensinya menggunakan PCA sebelum diproses oleh TFT yang dilatih dengan *physics-informed loss* dan dioptimasi menggunakan Optuna. Hasil prediksi TFT bersama data lingkungan membentuk keadaan berdimensi 33 bagi agen D3QN dalam memilih orientasi panel optimal. Evaluasi pada tujuh hari data uji menunjukkan model peramalan mencapai $R^2 = 0,9616$ dan $MAE = 0,239$ mW, sedangkan agen D3QN memperoleh EnergyRatio sebesar 0,9909 dan efisiensi 0,9769. Meskipun EnergyRatio D3QN lebih tinggi dibandingkan *baseline* panel statis (0,9882) dan *paired t-test* ($p = 0,175$) menunjukkan selisih tersebut belum signifikan secara statistik pada $n = 7$ hari. Studi ablasi turut mengungkap bahwa observasi daya aktual berkontribusi lebih besar dibandingkan prediksi TFT, bahkan varian tanpa TFT sedikit mengungguli model lengkap, meski integrasi TFT tetap mempercepat konvergensi pelatihan pada tahap awal. Temuan ini menegaskan bahwa kontribusi TFT lebih berperan sebagai panduan awal pembelajaran, bukan penentu performa akhir kebijakan agen.

Kata kunci: pembelajaran penguatan, *Double Dueling DQN*, *Temporal Fusion Transformer*, *physics-informed neural network*, orientasi panel surya

ABSTRACT

Solar energy is intermittent, requiring an adaptive panel-orientation control mechanism to maintain optimal energy yield. This study develops a solar panel orientation optimization system by integrating a Temporal Fusion Transformer (TFT) as the power forecasting model with a Double Dueling Deep Q-Network (D3QN) as the reinforcement learning agent. Primary data were collected from 13 fixed panel orientations and nine environmental variables over four weeks, dimensionally reduced via PCA, and processed by a TFT trained with a physics-informed loss and tuned using Optuna. The TFT predictions combined with environmental data form a 33-dimensional state for the D3QN agent to select the optimal panel orientation. Evaluation on seven days of test data shows the forecasting model achieving $R^2 = 0.9616$ and $MAE = 0.239$ mW, while the D3QN agent obtained an EnergyRatio of 0.9909 and efficiency of 0.9769. Although the D3QN EnergyRatio exceeded the static-panel baseline (0.9882) and paired t-test ($p = 0.175$) results indicate the difference was not statistically significant at $n = 7$ days. An ablation study further revealed that actual power observations contributed more than TFT predictions, with the no-TFT variant slightly outperforming the full model, although TFT integration still accelerated early training convergence. These findings indicate that TFT's contribution functions more as an early learning guide rather than a consistent determinant of the agent's final policy performance.

Keywords: reinforcement learning, Double Dueling DQN, Temporal Fusion Transformer, physics-informed neural network, solar panel orientation