

INTISARI

Reputasi yang baik sangat diharapkan oleh setiap Perguruan Tinggi (PT). Di Indonesia, status akreditasi PT menjadi tolok ukur reputasi baik suatu PT dan seringkali menjadi pertimbangan bagi calon mahasiswa yang mendaftar ke suatu PT. Penentuan status ‘Terakreditasi’ suatu PT didasarkan salah satunya pada capaian publikasi karya ilmiah dosen. Untuk memastikan perolehan status ‘Terakreditasi’ di periode selanjutnya, kemampuan untuk memprediksi capaian kinerja karya ilmiah menjadi aspek yang krusial dalam perencanaan strategis PT. Walaupun sebuah penelitian sudah mencoba menggunakan model deep learning LSTM untuk memprediksi jumlah publikasi individu, namun masih belum dapat menentukan model deep learning terbaik khususnya untuk memprediksi jumlah publikasi sebagai capaian kinerja.

Penelitian ini berfokus pada prediksi capaian kinerja publikasi karya ilmiah dosen suatu PT dengan menggunakan teknik pembelajaran mesin *deep learning*. Digunakan dan dibandingkan tiga model *deep learning* yang telah terbukti unggul untuk melakukan *forecasting* dari *sequential time series data*, yaitu *Long Short-Term Memory* (LSTM), *Bidirectional LSTM* (BiLSTM), dan *Convolutional LSTM* (ConvLSTM). Untuk meningkatkan akurasi dari model *deep learning* tersebut, dibandingkan *hyperparameter tuning* metode manual dengan metode *Bayesian optimization* pada setiap model *deep learning*.

Studi ini menggunakan data jumlah publikasi karya ilmiah suatu PT dari tahun 2010 hingga 2023 yang merupakan salah satu indikator kinerja penting dalam prediksi status akreditasi PT. Data ini didapatkan dari SINTA yang telah mengumpulkan data dari berbagai database publikasi. Data jumlah publikasi karya ilmiah digunakan untuk melatih model *deep learning* agar dapat memprediksi jumlah publikasi karya ilmiah suatu PT pada tahun selanjutnya. Model *deep learning* yang telah dilatih kemudian dievaluasi menggunakan *Mean Absolute Error* (MAE) untuk menentukan akurasi prediksi.

Hasil penelitian menunjukkan bahwa model LSTM dan BiLSTM memberikan prediksi paling akurat, dengan rata-rata MAE mencapai 294 pada LSTM dan 235 pada BiLSTM. Sementara itu, *Bayesian optimization* untuk *hyperparameter tuning* tidak secara signifikan meningkatkan kinerja model pada model BiLSTM dan ConvLSTM, bila dibandingkan dengan *hyperparameter tuning* secara manual. Berbeda dari BiLSTM dan ConvLSTM, penggunaan *Bayesian optimization* pada LSTM memberikan MAE yang lebih tinggi/lebih buruk dari *hyperparameter tuning* manual LSTM. Dari ketiga model yang ditawarkan dan metode *hyperparameter tuning* yang digunakan, BiLSTM dengan *Bayesian optimization* menjadi pilihan terbaik untuk memprediksi kinerja jumlah publikasi karya ilmiah suatu PT.

Kata kunci: Publikasi karya ilmiah, Perguruan Tinggi, *Deep learning*, LSTM, *Bayesian optimization*

ABSTRACT

A good reputation is highly sought after by every Higher education institution. In Indonesia, an institution's accreditation status serves as a benchmark for its reputation and is often a major consideration for prospective students when choosing a Higher education institution. One of the factors in determining an institution's "Accredited" status is the academic publication achievements of its faculty. To secure an "Accredited" status in the next period, the ability to predict future research publication performance becomes a crucial aspect of strategic planning for Higher education institution. Although a study has attempted to use deep learning LSTM models to predict individual publication counts, it still has not been able to determine the best deep learning model, particularly for predicting publication counts as a measure of performance achievement.

This study focuses on predicting the research publication performance of a Higher education institution faculty using deep learning machine learning techniques. Three deep learning models, known for their effectiveness in forecasting sequential time series data, are utilized and compared: Long Short-Term Memory (LSTM), Bidirectional LSTM (BiLSTM), and Convolutional LSTM (ConvLSTM). To enhance the accuracy of these deep learning models, this study compares manual hyperparameter tuning methods with Bayesian optimization for each model.

This study uses data on the number of research publications by a particular Higher education institution from 2010 to 2023, which serves as one of the key performance indicators in predicting accreditation status. This data was obtained from SINTA, which aggregates publication data from various databases. The publication data is used to train the deep learning models to predict the number of faculty research publications in the following year. The trained models are then evaluated using Mean Absolute Error (MAE) to determine the prediction accuracy.

The results indicate that the LSTM and BiLSTM models provide the most accurate predictions, with average MAEs of 294 for LSTM and 235 for BiLSTM. Meanwhile, Bayesian optimization for hyperparameter tuning does not significantly improve model performance for BiLSTM and ConvLSTM compared to manual hyperparameter tuning. In contrast to BiLSTM and ConvLSTM, using Bayesian optimization on LSTM results in higher/worse MAEs than manual tuning of LSTM. Among the three models and the hyperparameter tuning methods applied, BiLSTM with Bayesian optimization emerges as the best option for predicting the research publication performance of a Higher education institution.

Keywords: University Accreditation, Machine Learning, LSTM, Bayesian Optimization