

# Komparasi Kinerja Metode The Attention Mechanism-Enhanced LSTM (LSTM-ATT) dan Poisson Regression dalam Prediksi Insiden Demam Berdarah Dengue di DKI Jakarta dengan Mempertimbangkan Faktor Iklim = Comparison of the Performance of The Attention Mechanism-Enhanced LSTM and Poisson Regression Methods in Predicting Dengue Hemorrhagic Fever Incidence in DKI Jakarta by Considering Climate Factors

Hilmi Tito Shalahudin, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920551509&lokasi=lokal>

---

## Abstrak

Demam Berdarah Dengue (DBD) adalah penyakit yang disebabkan oleh virus dengue yang termasuk anggota genus Flavivirus dan keluarga Flaviviridae yang menyebar melalui nyamuk Aedes (Stegomyia). Jumlah kasus DBD di seluruh dunia dilaporkan meningkat lebih dari 8 kali lipat selama dua dekade terakhir, dari 505.430 kasus pada tahun 2000, menjadi lebih dari 2,4 juta pada tahun 2010, dan 5,2 juta pada tahun 2019. Peningkatan insiden DBD dipengaruhi oleh berbagai faktor, baik itu faktor alam, kebiasaan manusia, hingga jenis virus penyebab DBD itu sendiri. Faktor alam yang dimaksud disini antara lain faktor iklim, seperti curah hujan (mm), temperatur rata-rata ( $\hat{a}$ ), dan kelembapan rata-rata (%). Peningkatan insiden DBD dapat dicegah dengan upaya pencegahan yang dilakukan sedini mungkin oleh pemerintah dan masyarakat. Informasi prediksi tren insiden DBD dapat menjadi masukan bagi masyarakat dan pemerintah untuk meningkatkan kewaspadaan terhadap peningkatan insiden DBD di DKI Jakarta. Dalam melakukan prediksi tren insiden DBD, dapat dilakukan dengan berbagai pendekatan, salah satu diantaranya adalah machine learning. Pada tugas akhir ini, implementasi machine learning menggunakan model The Attention Mechanism-Enhanced LSTM (LSTM-ATT) dan Poisson Regression, akan digunakan untuk memprediksi tren insiden DBD dari waktu ke waktu. LSTM-ATT adalah sebuah model machine learning yang digunakan untuk memproses data sequence. Poisson Regression adalah model regresi yang dapat digunakan pada data yang variabel responnya berdistribusi Poisson dan bertipe diskrit. Prediksi yang akan dilakukan didasarkan pada jumlah insiden DBD sebagai variabel respon, serta faktor iklim seperti curah hujan, suhu, dan kelembapan sebagai variabel prediktor. Untuk proporsi data, kedua model tersebut menggunakan proporsi data training sebesar 80% dan data testing sebesar 20%. Model yang dibentuk ini dievaluasi dengan nilai dari Root Mean Square Error (RMSE) dan Mean Absolute Error (MAE). Hasil implementasi terbaik pada skripsi ini dihasilkan oleh model LSTM-ATT dengan evaluasi di setiap wilayah DKI Jakarta: Jakarta Pusat RMSE: 9,5727, MAE: 6,6946; Jakarta Timur RMSE: 21,5288, MAE: 15,6365; Jakarta Barat RMSE: 16,3683, MAE: 12,4908; Jakarta Utara RMSE: 23,5911, MAE: 15,2969; Jakarta Selatan RMSE: 18,3811, MAE: 14,0262.

.....Dengue Hemorrhagic Fever (DHF) is a disease caused by the dengue virus, which belongs to the Flavivirus genus and Flaviviridae family, transmitted through Aedes mosquitoes (Stegomyia). The number of DHF cases worldwide has reported an increase of more than 8 times over the past two decades, from 505,430 cases in 2000 to over 2.4 million in 2010 and 5.2 million in 2019. The increase in DHF incidence is influenced by various factors, including natural factors, human behavior, and the type of dengue virus itself. Natural factors include climate factors such as rainfall (mm), average temperature ( $\hat{a}$ ), and average humidity

(%). The increase in DHF incidence can be prevented through early prevention efforts by the government and the community. Predicting the trend of DHF incidence can provide input for the community and the government to increase vigilance against the increasing incidence of DHF in Jakarta. Various approaches can be used for predicting the trend of DHF incidence, one of which is machine learning. In this final project, the implementation of machine learning using the LSTM-ATT and Poisson Regression models will be used to predict the trend of DHF incidence over time. LSTM-ATT is a machine learning model used to process sequential data, such as time series data. On the other hand, Poisson Regression is a regression model that can be used for data with Poisson-distributed response variables and discrete types. The prediction will be based on the number of DHF incidents as the dependent variable, and climate factors such as rainfall, temperature, and humidity as independent variables. The performance of the models will be evaluated using the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). The best implementation results in this thesis were produced by the LSTM-ATT model with evaluation in each area of DKI Jakarta: Central Jakarta RMSE: 9.5727, MAE: 6.6946; East Jakarta RMSE: 21.5288, MAE: 15.6365; West Jakarta RMSE: 16.3683, MAE: 12.4908; North Jakarta RMSE: 23.5911, MAE: 15.2969; South Jakarta RMSE: 18.3811, MAE: 14.0262.