

ABSTRAK

Oka Nuantara, I Komang (2023), *Prediksi Waktu Tunggu Antrean Pada Mal Pelayanan Publik Dengan Metode XGBoost dan Random Forest*
Tesis, S2 Ilmu Komputer, Program Pascasarjana, Universitas Pendidikan Ganesha.

Tesis ini sudah disetujui dan diperiksa oleh Pembimbing I: Prof. Dr. Ir. Dewa Gede Hendra Divayana, S.Kom., M.Kom. dan Pembimbing II: Dr. Luh Joni Erawati Dewi, S.T., M.PD.

Kata Kunci: Machine Learning, XGBoost, Random Forest, Lagged Features, Forecasting

Penelitian ini bertujuan untuk membangun model prediksi waktu tunggu pada sistem antrean Mal Pelayanan Publik Kabupaten Badung (MPP Badung) menggunakan dua metode *Machine Learning*, yaitu Extreme Gradient Boosting (XGBoost) dan *Random Forest* (RF). Data waktu tunggu antrean dari dua loket layanan MPP Badung digunakan untuk membangun model prediksi, yaitu loket BPJS Kesehatan dengan model antrean tunggal server tunggal dan loket Dinas Kesehatan Kabupaten Badung (Dinkes Badung) dengan model antrean tunggal server ganda. Dataset yang dibangun juga akan digabungkan dengan data riwayat cuaca. Dataset yang dikumpulkan selanjutnya akan melalui tahap *pre-processing* seperti pembersihan, transformasi, integrasi, reduksi dan pembagian data. Dataset antrean akan dibagi menjadi tiga data, yaitu data latih untuk melatih model, data validasi untuk menguji performa model serta menyesuaikan hyperparameter dan data tes untuk menguji model menggunakan data yang belum pernah dilihat sebelumnya. Metode evaluasi yang digunakan pada tahap validasi dan tes menggunakan metrik MAE dan RMSE. Penelitian ini juga mengevaluasi apakah penambahan fitur riwayat antrean sebelumnya atau *lagged features* dapat meningkatkan akurasi model prediksi. Hasil penelitian menunjukkan bahwa kedua metode, XGBoost dan RF, mampu memprediksi waktu tunggu dengan tingkat akurasi yang tinggi pada saat pengujian, model XGBoost memberikan hasil yang lebih baik dibandingkan dengan RF. Pada dataset BPJS Kesehatan, model XGBoost mendapatkan nilai MAE 2.24 menit dan RMSE 3.17 menit sementara model RF memperoleh nilai MAE 2.61 menit dan RMSE 3.66. Selanjutnya, pada dataset Dinkes Badung model XGBoost memperoleh nilai MAE 4.47 menit dan RMSE 6.40 menit, sementara model RF mendapatkan nilai MAE 4.57 dan RMSE 6.37 menit. Penambahan fitur lagged atau data antrean sebelumnya secara konsisten memperbaiki kinerja model. Peningkatan akurasi model prediksi pada dataset BPJS Kesehatan dengan *lagged features* pada model XGBoost adalah sebesar 31.5% dan model RF sebesar 23.5%. Peningkatan akurasi model prediksi model Dinkes Badung dengan *lagged features* dengan XGBoost adalah sebesar 24.32% dan RF sebesar 23.35%. Fitur-fitur penting yang digunakan dalam prediksi waktu tunggu antrean meliputi jumlah sisa antrean, waktu tunggu rata-rata, waktu pembuatan antrian per menit, dan fitur-fitur terkait waktu dan pola antrean.

ABSTRACT

Oka Nuantara, I Komang (2023), *Queue Waiting Time Prediction in the Public Service Mall Using XGBoost and Random Forest Methods*.

Master's Thesis, Computer Science, Postgraduate Program, Ganesha University of Education.

This thesis has been approved and examined by Supervisor I: Prof. Dr. Ir. Dewa Gede Hendra Divayana, S.Kom., M.Kom. and Supervisor II: Dr. Luh Joni Erawati Dewi, S.T., M.PD.

Keywords: Machine Learning, XGBoost, Random Forest, Lagged Features, Forecasting

The purpose of this research is to build a predictive model for queue waiting times within the queue management system of the Public Service Mall of Badung Regency (MPP Badung) using two Machine Learning methods, namely Extreme Gradient Boosting (XGBoost) and Random Forest (RF). Queue waiting time data from two service counters at MPP Badung are utilized to build predictive models: the Social Security Agency in Health counter (BPJS Kesehatan) with a single-server single-queue model and the Badung Regency Health Office (Dinkes Badung) counter with a single-server multi-queue model. The constructed dataset is also merged with historical weather data. The collected dataset then undergoes pre-processing stages, including cleaning, transformation, integration, reduction, and data splitting. The queue dataset is divided into three subsets: training data for model training, validation data to test model performance and adjust hyperparameters, and test data to evaluate the model using previously unseen data. Evaluation methods during validation and testing employ the MAE and RMSE metrics. This study also evaluates whether the addition of lagged features can enhance the prediction model's accuracy. The results indicate that both XGBoost and RF methods effectively predict waiting times with high accuracy during testing, with XGBoost outperforming RF. In the BPJS Health dataset, XGBoost achieves an MAE of 2.24 minutes and an RMSE of 3.17 minutes, while RF records an MAE of approximately 2.61 minutes and an RMSE of around 3.66 minutes. In the Dinkes Badung dataset, XGBoost obtains an MAE of 4.47 minutes and an RMSE of 6.40 minutes, while RF receives an MAE of around 4.57 minutes and an RMSE of about 6.37 minutes. The consistent addition of lagged features consistently improves model performance. The improvement in prediction model accuracy on the BPJS Health dataset with lagged features for the XGBoost model is 31.5%, while for the RF model, it is 23.5%. The accuracy improvement in the Dinkes Badung model with lagged features is 24.32% for XGBoost and 23.35% for RF. Important features used in predicting queue wait times include the remaining number of queues, average wait time, queue creation time per minute, and features related to time and queue patterns.