

ABSTRAK

PENERAPAN *AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA)* UNTUK MEMPREDIKSI JUMLAH KENDARAAN MAL BERDASARKAN DATA PARKIR KENDARAAN

Oleh:

Andhika Eka Putra Sutrisno

1217050010

Meskipun Indonesia merupakan salah satu negara dengan tingkat kunjungan mal tertinggi, pola kunjungan tersebut belum banyak dipahami secara kuantitatif. Penelitian ini menerapkan metode *Autoregressive Integrated Moving Average (ARIMA)* untuk menganalisis dan memprediksi jumlah parkir kendaraan (roda empat) di mal berdasarkan data parkir kendaraan bulanan periode 2015–2025. Proses penelitian mengikuti kerangka kerja CRISP-DM yang terdiri atas pemahaman bisnis, pemahaman dan persiapan data, pemodelan, evaluasi, dan deployment. Pada tahap persiapan data, dilakukan deteksi dan penanganan anomali menggunakan *Modified Z-Score* dengan *threshold* 3,5 untuk mengidentifikasi penurunan ekstrem akibat pandemi COVID-19, yang kemudian diimputasi menggunakan *median rolling* dengan *window* 12 bulan. Data ditransformasi menggunakan logaritma natural untuk menstabilkan variansi, kemudian diuji stasionaritasnya menggunakan *Augmented Dickey-Fuller (ADF) Test*. Parameter model optimal ditentukan melalui *grid search* berbasis *Akaike Information Criterion (AIC)*, menghasilkan model ARIMA(2,1,3) sebagai parameter terbaik pada skenario pembagian data 80% *training* dan 20% *testing*. Hasil evaluasi pada data *testing* menunjukkan model mampu memberikan prediksi dengan *Mean Absolute Percentage Error (MAPE)* sebesar 11,88% yang berada pada kategori akurasi baik, *Mean Absolute Error (MAE)* sebesar 2.535 kendaraan per bulan, *Root Mean Squared Error (RMSE)* sebesar 3.108 kendaraan, dan *Normalized Mean Squared Error (nMSE)* sebesar 0,0194. Model kemudian dilatih ulang menggunakan seluruh data historis untuk menghasilkan prediksi jumlah kendaraan parkir 36 bulan ke depan disertai *confidence interval* 95%. Penelitian ini diimplementasikan dalam bentuk aplikasi *web* interaktif berbasis *Streamlit* yang memungkinkan eksplorasi data, konfigurasi metode deteksi anomali, dan visualisasi hasil prediksi secara langsung. Hasil penelitian ini diharapkan dapat memberikan wawasan bagi pengelola mal dalam merencanakan kapasitas parkir, alokasi sumber daya, dan strategi pemasaran yang lebih efektif berdasarkan proyeksi kunjungan di masa mendatang.

Kata Kunci: Kecerdasan Buatan, Pembelajaran Mesin, ARIMA, *Time Series*, Jumlah Parkir Kendaraan Mal, *statsmodels*.

ABSTRACT

APPLICATION OF AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) TO FORECAST THE NUMBER OF VEHICLES IN MALL PARKING BASED ON PARKING DATA

By:

Andhika Eka Putra Sutrisno

1217050010

Although Indonesia is one of the countries with the highest shopping mall visitation rates, these visitation patterns have not yet been well understood from a quantitative perspective. This study applies the Autoregressive Integrated Moving Average (ARIMA) method to analyze and predict the number of parked vehicles (four-wheeled) at shopping malls based on monthly vehicle parking data for the period 2015–2025. The research process follows the CRISP-DM framework, which consists of business understanding, data understanding and preparation, modeling, evaluation, and deployment. During the data preparation stage, anomalies were detected and handled using the Modified Z-Score with a threshold of 3.5 to identify extreme declines caused by the COVID-19 pandemic, which were then imputed using a rolling median with a 12-month window. The data were transformed using the natural logarithm to stabilize the variance, then tested for stationarity using the Augmented Dickey-Fuller (ADF) test. The optimal model parameters were determined through a grid search based on the Akaike Information Criterion (AIC), resulting in an ARIMA(2,1,3) model as the best fit for a data split of 80% training and 20% testing. Evaluation results on the test data show that the model is capable of providing predictions with a Mean Absolute Percentage Error (MAPE) of 11.88%, which falls into the “good accuracy” category; a Mean Absolute Error (MAE) of 2,535 vehicles per month; a Root Mean Squared Error (RMSE) of 3,108 vehicles; and a Normalized Mean Squared Error (nMSE) of 0,0194. The model was then retrained using all historical data to generate predictions of the number of parked vehicles for the next 36 months, along with a 95% confidence interval. This research was implemented as an interactive web application based on Streamlit, which allows for data exploration, configuration of anomaly detection methods, and real-time visualization of prediction results. The findings of this study are expected to provide insights for mall managers in planning parking capacity, resource allocation, and more effective marketing strategies based on future visit projections.

Keywords: ARIMA, machine learning, dan artificial intelligence, Time Series, Number of Parking Spaces at the Mall, statsmodels