



International Commission on Illumination
Commission Internationale de l'Éclairage
Internationale Beleuchtungskommission

MACHINE LEARNING BASED REAL-TIME PREDICTION OF LIGHTING CLASSES IN ADAPTIVE ROAD LIGHTING SYSTEMS

Onaygil, S., et al.

DOI 10.25039/x051.2025/kr3d33

This article is also published as part of:

Proceedings of the CIE 2025 Midterm Meeting Vienna, Austria, July 4-11, 2025:
Scientific Conference (July 7-9, 2025)

DOI 10.25039/x051.2025

in

Proceedings of the CIE (International Commission on Illumination)

ISSN no. 3061-015X (print), 3061-0168 (online)

The paper has undergone double-blind peer review and its final version has been presented at the CIE 2025 Midterm Meeting, Vienna, Austria, July 4–11, 2025.

© CIE 2025

All rights reserved. This work is licensed under the Creative Commons Attribution-NonCommercial 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>). Any mention of organizations or products does not imply endorsement by the CIE.

CIE Central Bureau
Babenbergerstrasse 9/9A
A-1010 Vienna, Austria
Tel.: +43 1 714 31 87
e-mail: ciecb@cie.co.at — www.cie.co.at

MACHINE LEARNING BASED REAL-TIME PREDICTION OF LIGHTING CLASSES IN ADAPTIVE ROAD LIGHTING SYSTEMS

Onaygil, S.¹, Yurtseven, M.B.¹, Feyzi, S.², Güler, Ö.¹, Erdoğan, S.², Oktay, G.²

¹ Istanbul Technical University, Energy Institute, Istanbul, TÜRKİYE; ² ISBAK Istanbul IT and Smart City Technologies INC., Istanbul, TÜRKİYE

onaygil@itu.edu.tr

Abstract

This study presents a machine learning-based framework for real-time prediction of road lighting classes in adaptive lighting systems. A pilot implementation was conducted on Büyükdere Street in Istanbul using two years of traffic and meteorological data. Lighting scenarios were developed in line with EN 13201 standards, and simulations with dimmable LED luminaires ensured compliance with lighting quality requirements. Hourly traffic flow and vehicle speed were predicted using Extreme Gradient Boosting and mapped to appropriate lighting classes via a rule-based matrix. The system was validated over six months, capturing seasonal and daily variations. The models achieved R^2 scores of 0.864 and 0.822 for flow and speed predictions, respectively. Predicted lighting classes reached 92.1% and 89.7% agreement with actual classes in two directions. Compared to non-adaptive M2-level lighting, the proposed system provided up to 32.1% energy savings. Results confirm the feasibility of data-driven adaptive lighting for energy-efficient and safe urban road environments.

Keywords: Energy Efficiency, Road Lighting Class, Adaptive Road Lighting, Control and Automation, Road Lighting Automation Algorithm, LED, Machine Learning

1 Introduction

Efficient road lighting plays a vital role in ensuring driver safety while minimizing energy consumption in urban environments. While traditional lighting systems may be insufficient in providing energy efficiency, adaptive road lighting systems aim to solve this problem by utilising real-time data. However, determining automation strategies and scenarios that can be applied in adaptive road lighting, where there are many variable parameters and adaptation times to different scenarios, is critical due to high-speed movement, and is still an open research topic. Adaptive road lighting algorithms, which are created according to certain assumptions and adapted to road lighting installations, also need to be analysed and approved for their suitability in real conditions. In this study, Büyükdere Street in the Sarıyer district of Istanbul was selected as a pilot road for smart lighting application, and a comprehensive analysis was performed on adaptive lighting systems.

The pilot road was evaluated following EN 13201 standards to determine the lighting class according to traffic density and speed scenarios. Within the scope of the study, the performances of the LED luminaires used at different dimming levels were measured, and lighting simulations were performed to ensure energy efficiency and lighting quality criteria. A comprehensive set of traffic, meteorological, temporal, and contextual features was used to estimate hourly flow rate and average vehicle speed via the XGBoost algorithm (Chen and Guestrin, 2016). Based on these predictions, suitable road lighting classes were assigned under predefined scenarios. The determined road lighting automation algorithm was defined in the pilot road system and the installation was implemented. The compatibility of the real lighting classes calculated from the real traffic and speed measurements on the pilot road was compared with the lighting classes estimated based on historical data.

2 Pilot Road and Road Lighting Simulations

The pilot road has six lanes, three going and three returning, and a 3.5-meter-wide median strip between the two directions. In the current situation, LED luminaires are mounted on 2-meter-

long consoles on 13-meter-high poles with a 40-meter spacing, and a twin-bracket central lighting arrangement is used.

The pilot road's lighting class was determined as M2 using the CEN/TR 13201-1: 2014, "Road lighting – Part 1: Guidelines on selection on lighting classes" (CEN/TR, 2014). Considering EN 13201-5: 2015, "Road Lighting – Energy Performance Indicators", which allows lowering the lighting class by up to two levels for driver comfort, the minimum lighting class was set as M4 (EN, 2015). Simulations for M2, M3, and M4 classes were conducted.

To determine the operating power levels of LED luminaires that meet the lighting requirements for M2, M3, and M4 classes on the pilot road, electrical and photometric measurements were carried out at the ITU Energy Efficiency and Lighting Technology Laboratory using an ENTEK MPR-60S Energy Analyser and a Labsphere 2-meter Ulbricht Sphere. Simulations ensured the luminaires met the luminous flux requirements of M2, M3, and M4 lighting classes. Results are presented in Table 1.

Table 1. Simulation Results vs. Required Values

	L_{ave} (cd/m ²)	U_0	U_l	f_{TI} (%)	R_{EI}
Required Values (M2)	1.50	0.40	0.7	10	0.50
Simulation Results (M2)	1.50	0.65	0.84	8	0.92
Required Values (M3)	1.00	0.40	0.50	15	0.50
Simulation Results (M3)	1.02	0.65	0.84	7	0.92
Required Values (M4)	0.75	0.40	0.60	7	0.92
Simulation Results (M4)	0.75	0.65	0.84	7	0.92

* L_{ave} : Average Luminance (cd/m²), U_0 : Overall Uniformity, U_l : Longitudinal Uniformity, f_{TI} : Threshold Increment, R_{EI} : Edge Illuminance Ratio

According to the simulation results, Table 2 gives the power and luminous flux values of the LED luminaires on the pilot road, which can provide lighting quality criteria for the M2-M3-M4 road lighting classes.

Table 2. LED Luminaire Power and Luminous Flux Values for Different Lighting Classes

Lighting Class	Power (W)	Luminous Flux (lm)
M2	134.7	19200
M3	86.7	13020
M4	63.1	9600

Based on the simulation results, the installation was carried out on the pilot road, and the luminaires' luminous flux levels were adjusted to obtain the M2, M3, and M4 road lighting classes.

3 Adaptive Road Lighting Automation System and Scenarios

Considering the traffic flow rate, average vehicle speed, and navigational task parameters, the pilot road lighting class was evaluated according to different combinations of these variables and the possible road lighting classes given in Table 3 were determined in line with the technical report CEN/TR 13201-1:2014 (CEN/TR, 2014).

Table 3. Pilot Road Adaptive Road Lighting Algorithm

		Traffic Flow Rate (vehicle/h)		
		$x > 65\%$ Max. Capacity	$35\% < x < 65\%$ Max. Capacity	$x < 35\%$ Max. Capacity
Average Vehicle Speed (km/h)	$70 < v < 100$ km/h	M2	M3	M4
	$40 < v \leq 70$ km/h	M3	M4	M4
	$v \leq 40$ km/h	M3	M4	M4

Within the scope of the study, data on the number of vehicles and vehicle speeds in each lane were obtained from the existing sensors on the road for two years back, at a frequency of two minutes. The maximum of the 2-minute flow rate data collected from three lanes in the same direction was selected and aggregated as the "hourly traffic flow rate." Similarly, the maximum speed value in the three lanes was accepted as the speed in that direction.

During the hours when the road lighting system is on, the highest traffic flow rate data defined for each lane and both directions is determined as the "maximum capacity". When the hourly data set is examined, it is seen that although the speed limit of the road is defined as 70 km/h, the average vehicle speed is above this value when the traffic on the road decreases. During the hours when the road is in heavy use, the average vehicle speed can be well below 70 km/h. With the help of this data analysis, road lighting classes were estimated using machine learning algorithms.

4 Data Preparation

A comprehensive data preparation workflow was implemented to ensure accurate predictions and alignment with real-world conditions. The dataset covers six months from July to December 2024, providing a diverse temporal span that includes seasonal and weekly traffic variations. This stage involved collecting heterogeneous data from multiple sources, cleaning and transforming them, and formatting them for machine learning modelling. Traffic data were collected at 2-minute intervals via RTMS Bluetooth-based sensors installed on Büyükdere Street in Istanbul. These included total vehicle count (flow rate) and average vehicle speed. To enrich the dataset:

- Meteorological data (temperature, precipitation) were integrated via the Copernicus API (Copernicus Data Space Ecosystem, n.d.), part of the EU Earth Observation Program.
- Official holidays, school breaks, and day-of-week features were obtained from national calendar APIs.
- Traffic announcements (accidents, events, parking) from the Urban Transport Management Centre were extracted but later excluded based on feature selection due to their outlier behaviour and low predictability.

Since raw traffic data was captured every two minutes, hourly summaries were generated by aggregating 30 consecutive measurements:

- Hourly vehicle count: The Maximum value across three lanes in the same direction is aggregated.
- Hourly average speed: The Maximum average speed among the three lanes, followed by hourly averaging, is aggregated.

To enhance temporal awareness, additional features were engineered from timestamps:

- Day, month, year, hour
- Day of week (Monday=0, ..., Sunday=6)
- Weekday vs weekend indicator (binary categorical)
- Week of year (to capture seasonal cycles)
- Seasonality and trend components: Generated with a period of 168 hours (7 days × 24 hours), to reflect weekly cycles in traffic dynamics.

Gaps in the 2-minute intervals were addressed via interpolation, particularly critical for flow rates sensitive to time resolution. Data types were standardised, and outliers and inconsistent entries were cleaned through validation and conversion steps.

Using min-max scaling, numerical variables such as temperature, precipitation, and hour were normalised to the [0, 1] range. Categorical variables (weekday/weekend, hour segments) were converted using one-hot encoding. In addition, the evening and nighttime hours, where lighting behaviour is more dynamic, were segmented into three categories (19:00–24:00, 00:00–02:00, and 02:00–06:00) and encoded as discrete input variables. The prepared data set was given as input to the machine learning algorithm. The overall workflow from data preparation to lighting class estimation is summarised in Figure 1.

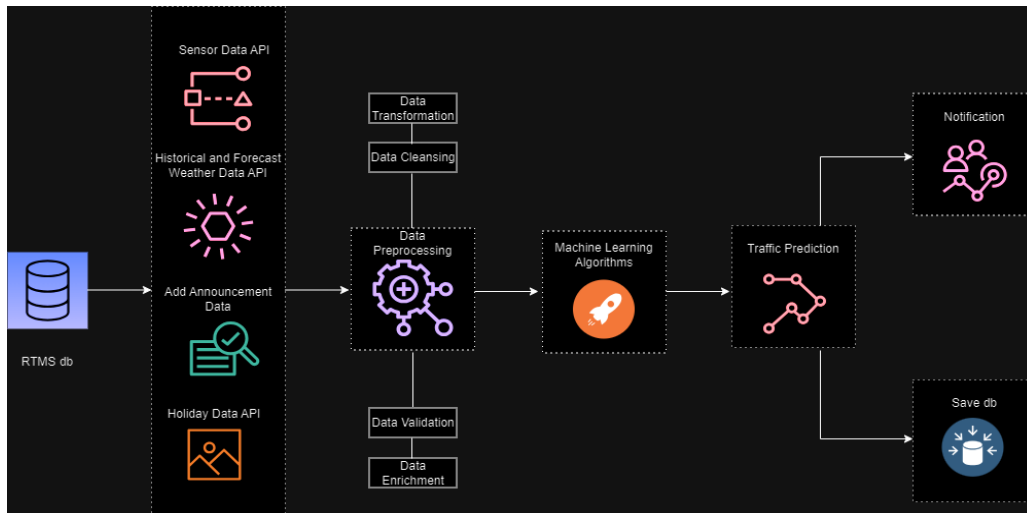


Figure 1. Road Lighting Class Estimation Workflow

In the next section, the data modelling process is detailed.

5 Modelling and Prediction of Traffic and Speed

Based on enriched historical datasets, the modelling phase aimed to predict hourly traffic flow rate and average vehicle speed using machine learning techniques. These predicted variables were later used to determine road lighting classes according to predefined rules (Table 3). Two separate regression models were trained:

- Model 1: Prediction of traffic flow rate (vehicle count per hour)
- Model 2: Prediction of average vehicle speed (km/h)

Although multi-output regression strategies were initially explored, it was found that training two independent models yielded superior performance and model interpretability. The models were implemented using the XGBoost algorithm, a highly efficient and scalable ensemble method based on decision trees. XGBoost builds models stage-wise, where each new tree corrects the residual errors of the previous ones. Its key strengths include:

- Handling missing values internally during training
- Built-in regularization to avoid overfitting
- Fast training speed and parallelisation
- Effective with both numerical and categorical inputs (after encoding)

Mathematically, the boosting process follows Equation 1:

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \quad (1)$$

Where:

$F_m(x)$: prediction at iteration m

$F_{m-1}(x)$: prediction at the previous iteration ($m-1$)

γ_m : learning rate controlling the contribution of each tree

$h_m(x)$: newly fitted regression tree at iteration m

This formulation allows the model to iteratively improve its predictions by correcting previous residual errors. The process continues for m steps, or until convergence, each time adding a new tree weighted by the learning rate γ_m .

In this study, the initial model $F_0(x)$ was set as the mean of the target variable, and subsequent iterations added trees to reduce the mean squared error (MSE) between predictions and actual values.

For hyperparameter tuning, Grid Search was employed to identify the optimal values for:

- Learning rate
- Max. depth
- Min. child weight
- N estimators

These were selected based on cross-validation scores using the MSE metric as the primary evaluation criterion. Once optimal parameters were identified, the final models were trained and serialised for deployment.

6 Estimated and Actual Road Lighting Classes

Following the hourly traffic flow rate and average vehicle speed prediction, the appropriate road lighting class for each hour was determined using a rule-based classification matrix (Table 3). This matrix was derived from EN 13201 standards and adapted for the pilot road segment. It maps combinations of vehicle count and average speed to lighting classes M2, M3, and M4. This decision structure was encoded as a lookup function and applied to each predicted flow and speed hourly pair. The resulting lighting class was logged alongside the timestamp and system direction. By leveraging this approach, the lighting system dynamically adjusts luminance levels to match real-world traffic conditions, ensuring energy savings and safety. The final predicted lighting classes were then compared to actual classes recorded by the system to assess performance.

The performance of the prediction models was evaluated using standard regression error metrics. For traffic flow rate prediction, the model achieved a Root Mean Squared Error (RMSE) of approximately 13.16 vehicles/hour, a Mean Absolute Error (MAE) of 8.25 vehicles/hour, and an R^2 score of 0.864, indicating a strong correlation between predicted and actual flow values. Similarly, for the average vehicle speed prediction, the RMSE was 6.54 km/h, the MAE was 3.73 km/h, and the R^2 score reached 0.822. These results confirm that the XGBoost-based models provided reliable estimates of both traffic volume and speed, forming a robust basis for accurate lighting class decisions. The high R^2 values indicate that the models effectively captured the underlying patterns and variations in the traffic data under diverse conditions.

To assess the accuracy of the lighting class assignment, the predicted lighting classes derived from the machine learning model were compared with the reference classes that were calculated from real traffic and speed measurements using the same rule-based matrix (Table 3). This comparison yielded high agreement levels, with an accuracy score of 92.1% for Direction 1 and 89.7% for Direction 2. These results demonstrate that the machine learning approach approximates the rule-based decisions derived from actual data, confirming its effectiveness for real-time adaptive lighting classification.

Figure 2 presents the normalised monthly distribution of real lighting classes derived from actual traffic and speed. Figure 3 displays the corresponding distribution of predicted lighting classes generated by the machine learning model. Class M3 consistently dominates across all months in both plots, reflecting typical mid-level traffic conditions on the pilot road. The predicted distribution closely mirrors the real data, particularly in the proportional use of M3 and M4 classes. Minor discrepancies, such as a slight overestimation of Class M2 in July, August and October, suggest that the model slightly over-predicts higher illumination levels under certain flow-speed combinations. The visual comparison confirms the model's capability to generalise well across seasonal and monthly traffic variations.

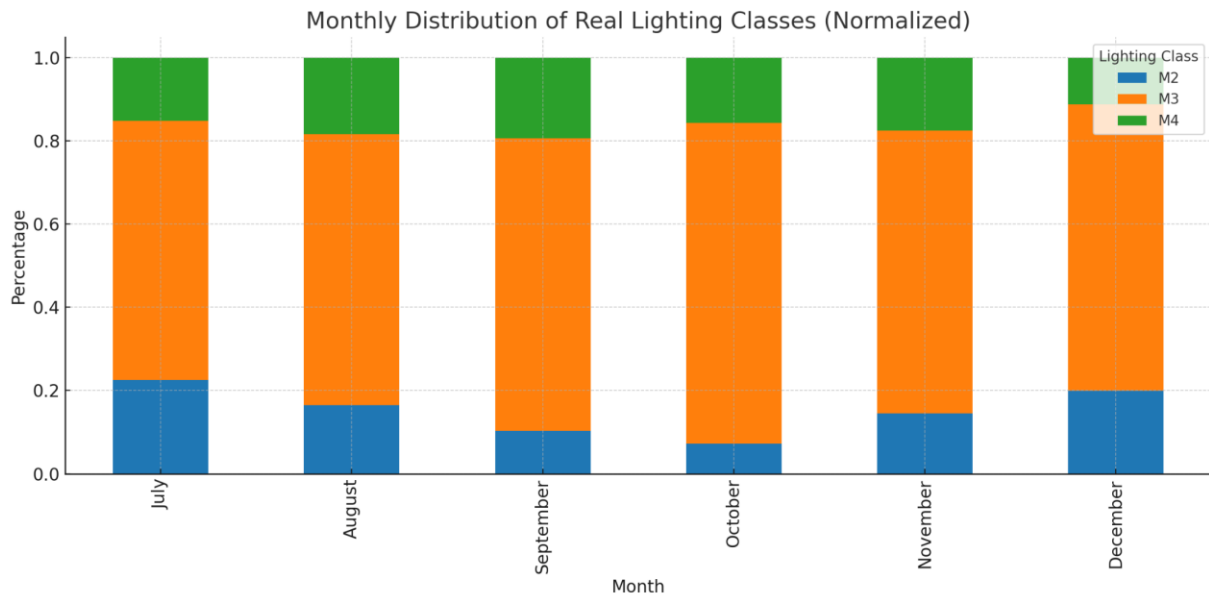


Figure 2. Normalised Distribution of Monthly Lighting Classes for Real Sensor Data

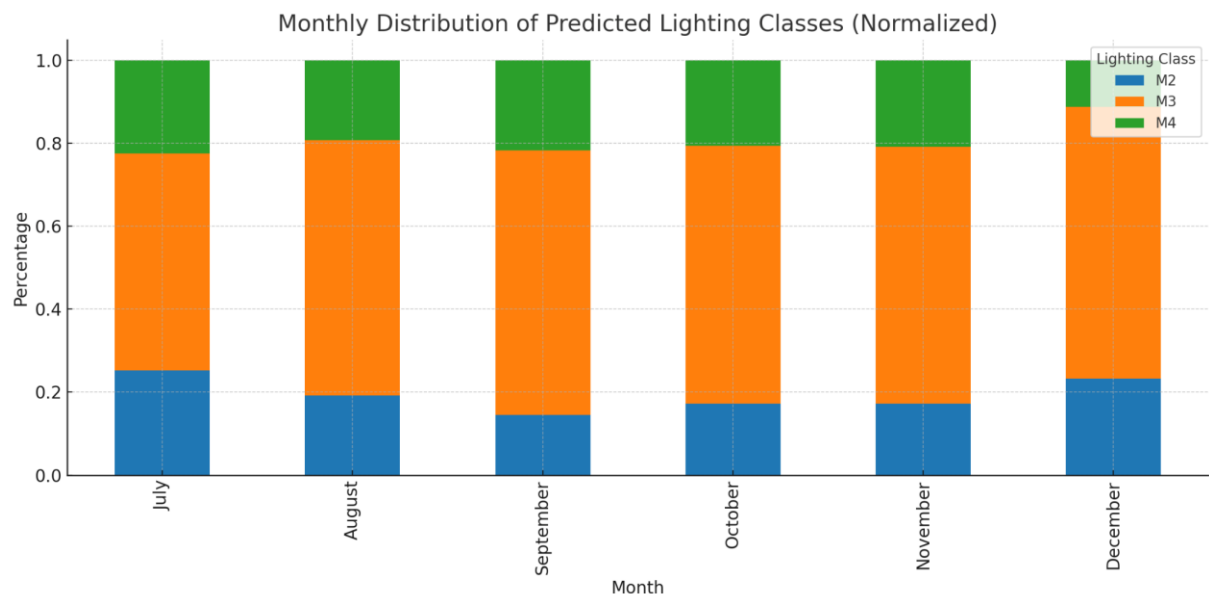


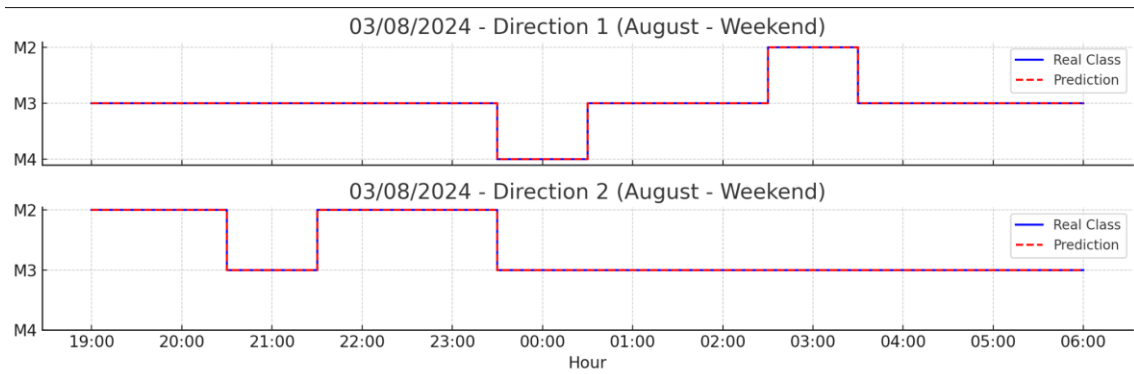
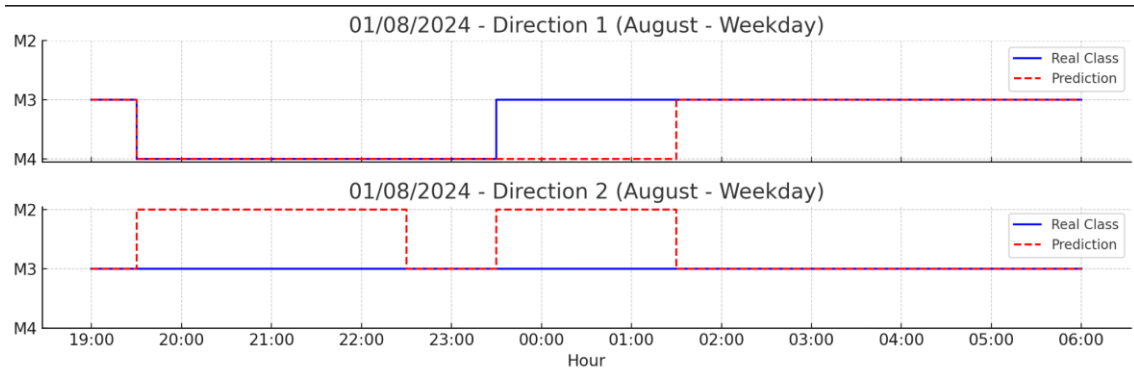
Figure 3. Normalised Distribution of Monthly Lighting Classes for Predicted Data

Six representative days were selected to represent different seasons across three different months to provide a detailed day-level evaluation of the model's performance, covering both weekdays and weekends. The selected dates include:

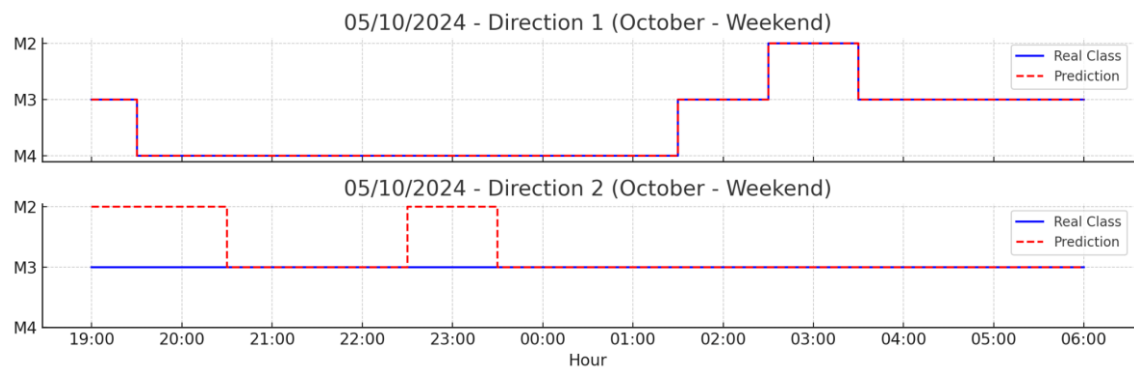
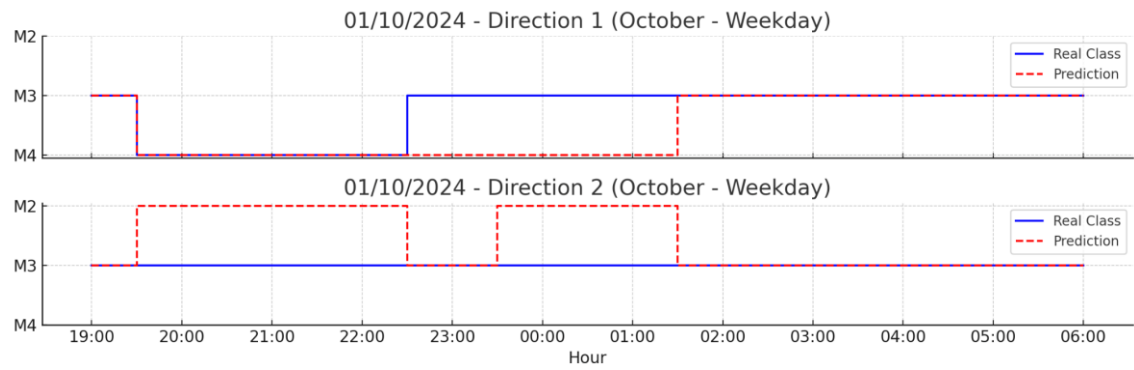
- 01.08.2024 (weekday) and 03.08.2024 (weekend) for August,
- 01.10.2024 (weekday) and 05.10.2024 (weekend) for October,
- 07.12.2024 (weekday) and 09.12.2024 (weekend) for December.

For each day, hourly lighting classes were derived from the actual traffic measurements (using the rule-based matrix) and the predictions generated by the machine learning model. The comparison is visualized in Figure 4, where the predicted lighting classes (dashed red lines) are plotted alongside the reference classes (solid blue lines) for both directions. These plots highlight the model's ability to closely follow real-world lighting decisions throughout the night hours, especially during transitional periods between lighting levels.

August



October



December

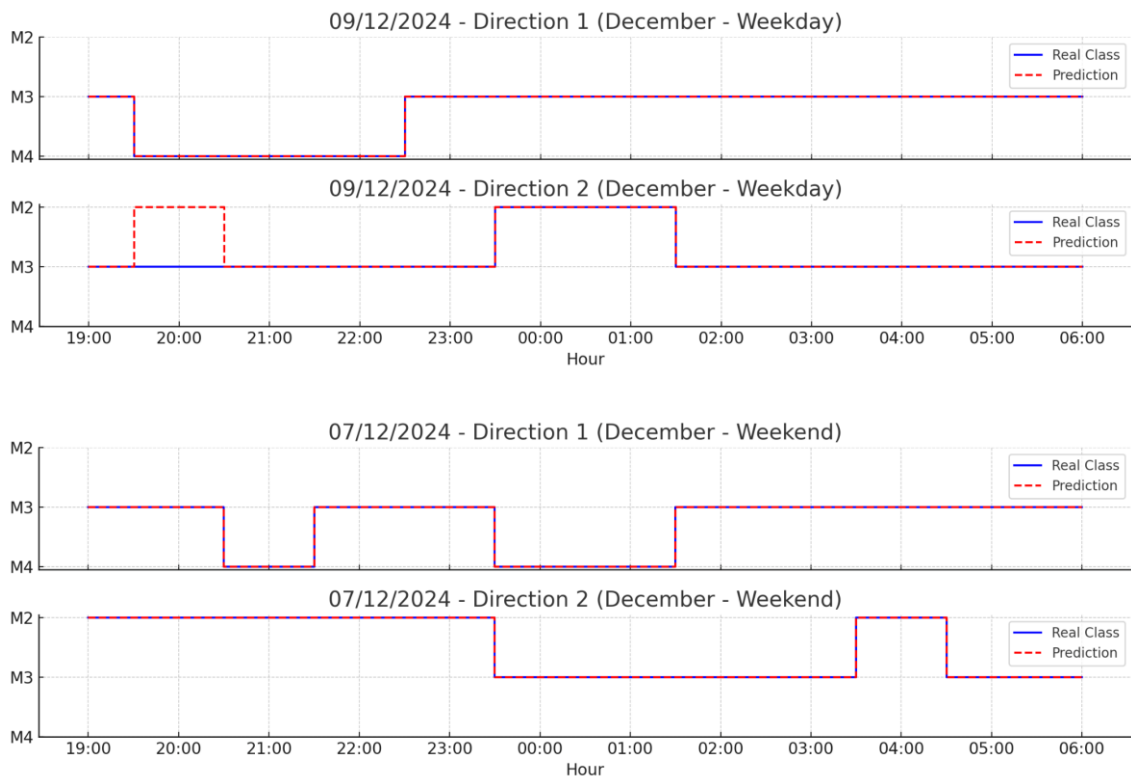


Figure 4. Hourly Lighting Class Comparison for Selected Weekdays and Weekends

To complement this day-level evaluation, Figure 5 summarises the model’s directional error tendencies monthly. Each bar represents the percentage of prediction errors that fall into each category. Overestimations occur when the model predicts a higher lighting class than required, while underestimations occur when the predicted class is lower than the actual class. This visualization allows for a better understanding of the model’s directional bias in different months and supports the evaluation of seasonal performance consistency.

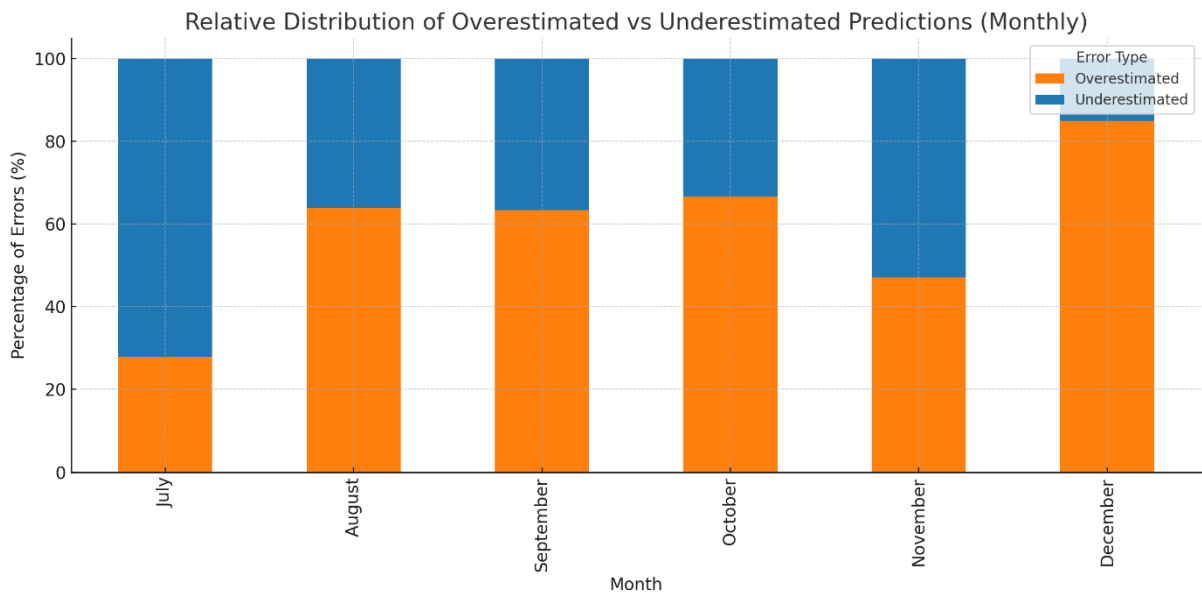


Figure 5. Relative Frequency of Overestimated and Underestimated Classes

While overestimations may result in unnecessary energy consumption, underestimations present a more critical concern as they indicate situations where the lighting class was predicted lower than required. This can lead to insufficient illumination and may compromise road safety. Therefore, understanding the frequency and distribution of underestimations is essential for evaluating the model's operational reliability and safety performance.

7 Energy Consumption and Efficiency Evaluation

To evaluate the impact of lighting class predictions on energy usage, a relative comparison was made using three lighting scenarios as given in Table 2:

- M2 class operation, where each luminaire consumes 134.7W,
- M3 class operation at 86.7W,
- M4 class operation at 63.1W.

Assuming a non-adaptive system where luminaires would operate continuously at the M2 class level (the highest energy-consuming scenario), monthly energy consumption was calculated and compared with two adaptive alternatives:

- Based on the real lighting classes observed,
- Based on the classes predicted by the machine learning model.

Table 4 summarises the corresponding relative savings percentages for each case.

Table 4. Relative Energy Savings for Real and Predicted Classes

Month	Saving Ratio with Real Classes (%)	Saving Ratio with Predicted Classes (%)
July	30.28	30.56
August	32.99	32.19
September	35.34	34.26
October	35.81	33.12
November	33.54	33.17
December	30.50	29.33
Average	33.10	32.10

8 Results and Discussion

The findings of this study demonstrate the effectiveness of the machine learning-based approach in accurately predicting road lighting classes under real-world conditions. Over a six-month evaluation period (July–December 2024), the model strongly aligned with actual lighting decisions made by the rule-based matrix derived from real traffic and speed data.

The performance of the ML model in predicting traffic flow and speed, which are critical inputs for lighting class determination, was validated using standard metrics. The model achieved an R² of 0.864 for vehicle count and 0.822 for average speed, indicating robust predictive capacity. These estimates were then mapped to lighting classes using a predefined rule matrix, resulting in a high agreement between predicted and actual classes — 92.1% accuracy for Direction 1 and 89.7% for Direction 2.

Beyond accuracy, a monthly classification error analysis evaluated the model's predictive bias (Figure 5). While the rate of overestimated classes was generally higher than that of the underestimated ones, the latter is of greater concern from a safety perspective. These results suggest that the model tends to lean towards the conservative side, favouring higher lighting classes that ensure visibility but may lead to energy inefficiencies.

Energy savings were assessed by comparing predicted and real lighting class allocations against a non-adaptive baseline where luminaires operate at M2 levels continuously. As shown

in Table 4, the adaptive strategy based on real classes yielded an average energy saving of 33.1%, while the machine learning-based prediction scenario achieved 32.1% savings. This close alignment further validates the operational reliability of the proposed system in achieving energy efficiency goals without compromising safety.

This study demonstrates that machine learning can be effectively integrated into adaptive road lighting systems to optimise energy use and operational safety. The model achieved high predictive accuracy in estimating traffic flow and speed, leading to around 90% agreement with actual lighting decisions and enabling around 30% energy savings compared to non-adaptive lighting. Its six-month performance suggests strong resilience to seasonal and daily fluctuations, supporting its use in dynamic urban environments. Importantly, the framework is modular and adaptable, allowing easy retraining with local datasets and integration into existing sensor infrastructures. As cities increasingly prioritize sustainable and data-driven solutions, this approach offers a scalable and practical tool for intelligent lighting management in future-ready urban infrastructure.

References

- CEN, 2014. CEN/TR 13201-1: Road lighting – Part 1: Guidelines on selection on lighting classes. Brussels: European Committee for Standardization.
- Chen, T. and Guestrin, C., 2016. XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). New York: ACM, pp.785–794. doi:10.1145/2939672.2939785
- Copernicus Data Space Ecosystem, n.d. Copernicus Data Space Ecosystem. [online] Available at: <https://dataspace.copernicus.eu/> [Accessed 30 Apr. 2025].
- EN, 2015. EN 13201-5: Road lighting – Energy performance indicators. Brussels: European Committee for Standardization.