DOI: 10.2478/ijasitels-2025-0010

IoT System for Milk Dispenser Monitoring and Management

Claudia Banciu, Teodora-Artemisa Menchiu, Adrian Florea

Department of Computer Science and Electrical Engineering, "Lucian Blaga" University of Sibiu, Sibiu, Romania

Corresponding author: Claudia Banciu, e-mail: claudia.banciu@ulbsibiu.ro
Teodora Menchiu, e-mail: teodora.menchiu@ulbsibiu.ro
Adrian Florea, e-mail: adrian.florea@ulbsibiu.ro

Abstract

An Internet of Things (IoT)-like application is an advanced system for automation and data analytics that leverages communication networks, sensors, and artificial intelligence to provide a comprehensive solution for a product or service. The development of IoT technologies has facilitated advancements in automated systems across various industries like healthcare, agriculture, smart cities, smart homes, and transportation. This paper presents an IoT-based system designed for the real-time monitoring and management of a milk dispenser. The system integrates an ESP32 microcontroller, sensors for temperature and milk level measurement, and cloud storage for data analysis and visualization. By using machine learning algorithms, the system predicts milk consumption trends and detects potential dispenser blockages, enhancing operational efficiency. Experimental evaluation showed that the Random Forest and Gradient Boosting regression models achieved high predictive accuracy, with the best configurations reaching a coefficient of determination (R2) of 0.95, a Mean Absolute Error (MAE) as low as 1.05 liters, and a Root Mean Squared Error (RMSE) of 1.32 liters, confirming the strong reliability of the proposed approach.

Keywords: IoT system, milk-dispenser, machine learning algorithms

1 Introduction

Milk dispensers are specialized machines that are designed to store and dispense milk efficiently while maintaining optimal freshness. Despite their practicality, there are also challenges in managing milk dispensers. Some challenges are ensuring consistent milk availability, maintaining the correct storage temperature, avoiding blockages or any other failures.

Beyond their basic functionality, milk dispensers bring several practical and economic benefits that contribute to their growing adoption. First, they significantly reduce packaging waste, eliminating the need for single-use cartons or bottles, which makes them a more environmentally sustainable solution. Second, dispensers support cost efficiency, since purchasing milk in bulk is generally cheaper for institutions and businesses compared to buying multiple individual containers. Third, they increase convenience by providing quick, portion-controlled access to milk, which is particularly valuable in high-consumption environments such cafeterias, farmers' markets or grocery stores. Dispensers ensure freshness and hygiene, as they maintain milk at a safe chilled temperature and reduce the risk of contamination that could occur when repeatedly opening bottles or cartons.

1.1. Challenges in Milk Dispenser Adoption

Despite their advantages, milk dispensers also face several challenges that can limit widespread adoption. First, the initial cost of purchasing and installing the equipment can be relatively high. However, in our study, this limitation is less significant, as the system is already installed and provides real operational data for analysis. Second, dispensers require regular maintenance and cleaning to prevent bacterial growth and ensure hygiene standards, which can be time-consuming. Third, they involve continuous electricity use, since the milk must remain chilled at a constant temperature, resulting in additional energy costs. Finally, dispensers have limited portability compared to bottles or cartons, as their size and bulk make them unsuitable for milk transport outside of fixed locations.

1.2. Market Context and Adoption Trends

The adoption of milk dispensers is steadily increasing both in Europe and worldwide. According to [1] the global market for milk vending machines was valued at approximately USD 0.9 billion in 2023 and is projected to nearly double to USD 1.8 billion by 2032, with a compound annual growth rate (CAGR) of around 8.1%. Europe represents the largest regional market, driven by consumer demand for fresh milk and strong traditions in dairy farming. Countries such as Germany, Italy, and France already report widespread use of milk dispensers in supermarkets, schools, and farmers' markets.
In Romania, milk dispensers are also present, though their adoption is more recent. The Romanian company FERMAT has reported the sale of more than 1000 milk vending machines across the country. Furthermore, consumer studies indicate that Romanian buyers associate these machines with product freshness, sustainability (through reusable bottles), and direct access to local dairy producers.

1.3. IoT and milk dispensers

The IoT presents new possibilities for automating and optimizing these systems. An IoT system typically consists of three essential layers: the sensor layer (physical objects), the connectivity layer (networks such as Wi-Fi), and the cloud layer (data storage and algorithms). IoT enables the integration of sensors and microcontrollers to monitor milk levels and temperature in real-time, offering predictive tools for maintenance. The proposed system employs an ESP32 microcontroller, temperature and weight sensors, and cloud-based data analytics to improve operational efficiency and enhance user experience. The appearance of the IoT has opened new possibilities for optimizing and automating these devices. By leveraging IoT, sensors and microcontrollers can track milk levels and temperatures, while cloud platforms provide users with real-time data access. This connectivity aids in predictive maintenance, reduces downtime, and ensures a smooth user experience.

Through smart devices and data analysis, IoT makes it possible to monitor milk levels, track temperature, and even predict issue. Incorporating IoT into milk dispensers provides several significant advantages. For instance, real-time monitoring of milk levels and temperature ensures that the milk remains fresh and available when needed. Cloud platforms offer centralized data storage and analysis, enabling users to remotely access data through intuitive interfaces. Additionally, IoT systems can use machine learning algorithms to predict milk consumption patterns and detect potential issues, such as blockages, before they disrupt service.

This paper aims to present an IoT-based solution for milk dispenser management. The proposed system integrates an ESP32 microcontroller, temperature and weight sensors, and cloud-based analytics to offer real-time insights. By addressing common challenges and introducing predictive capabilities, the system seeks to improve the efficiency, reliability, and usability of milk dispensers.

The rest of this paper is structured as follows. Section 2 reviews the state of the art and related work in IoT-enabled applications targeting different fields such as agriculture, food industry and dairy supply chain. Section 3 describes the methodology used, presenting the hardware and software architecture of the IoT system. Section 4 presents and compares results obtained by different prediction techniques. Section 5 highlights the conclusions and presents future research directions.

2 Related Work

Over recent times, the adoption of this kind of technologies has become increasingly prevalent in agriculture and food monitoring systems, and many other fields like healthcare, smart homes, and even smart cities. There are numerous studies which emphasize the critical role that the intelligent monitoring and automation have in providing efficiency, sustainability, also enabling the real-time decision-making.

In [2] the authors provide a comprehensive review of Agriculture 4.0, highlighting the importance of integrating the IoT, the artificial intelligence, and automation in modern farming practices. The authors highlight several key benefits including optimizing the resource management, increasing the crop yields, monitoring the soil and weather in real time or even reducing the costs of labor.

Similarly, the authors of [3] demonstrate how IoT-enabled smart agriculture platforms improve crop yield and environmental monitoring through the deployment of sensor networks. In this study the authors discussed about the importance of using the WSNs (wireless sensor networks) which can advise the farmers about different aspects like the status of the agricultural machinery and about all the factors that can influence the crop growth. They also highlighted the significance of using machine learning algorithms which can help to detect or even prevent plant diseases. These techniques can be used to analyze the collected data and identify the initial signs of pest infestations or other diseases, which will allow for further effective actions.

In [4], the authors have developed an IoT platform using LoRaWAN (‘Long-range wide-area network’) technology specifically designed for smart irrigation. The study implements LoRaWAN within an IoT framework and it is used for precision irrigation. This system was tested in a plasticulture field for fresh-market tomato production and the obtained results indicate the strong potential in aspects of precision irrigation, yield improvement and water conservation.

Beyond the agricultural context, IoT and AI technologies are increasingly being adopted in the food industry and dairy supply chain, where they contribute to real-time monitoring, predictive analytics, and quality assurance. Recent studies demonstrate that IoT-based systems integrated with machine learning models can significantly enhance milk production management and safety. For example, [5] introduced an inexpensive AI-powered IoT sensor for continuous farm-to-factory milk quality monitoring. This system combines spectral analysis with machine learning models such as Random Forest to estimate protein and fat content, achieving an impressive accuracy of ±0.14% for fat and ±0.07% for protein. Similarly, [6] presented predictive modeling for milk production using artificial intelligence and machine learning techniques, emphasizing the ability of these models to capture complex nonlinear relationships between environmental, biological, and management factors that influence milk yield. The study compared several approaches, including Random Forest, Gradient Boosting, Support Vector Machines, and Artificial Neural Networks, demonstrating that ensemble and deep learning models consistently provided more reliable and accurate predictions than traditional statistical methods. The authors concluded that integrating AI and ML within dairy management systems can significantly improve forecasting capabilities, resource allocation, and decision-making efficiency in milk production.

In addition, [7] developed an IoT-based monitoring framework designed to preserve the quality and safety of dairy products throughout the cold supply chain. The proposed system, named i-EAT, integrates sensor-equipped crates (i-EAT Tags) and vehicle-level modules (i-EAT Trucks) capable of collecting temperature, humidity, vibration, pressure, and GPS location data during transportation. Using Bluetooth Low Energy (BLE) for local communication and cloud-based data storage, the system allows real-time tracking and early detection of anomalies, such as temperature fluctuations or excessive vibration during loading and delivery. The study demonstrated that this IoT infrastructure significantly enhances traceability, reduces waste, and increases consumer confidence by ensuring product quality and freshness from producer to retailer.

Complementing these efforts, [8] presented a digital IoT-enabled platform for milk collection and quality assessment using machine learning algorithms to detect adulteration in milk. Their model achieved an accuracy rate above 98%, demonstrating that ML-based IoT systems can effectively ensure product integrity and transparency across the dairy value chain.

Taken together, these studies provide strong evidence that combining IoT infrastructures with advanced machine learning algorithms can lead to substantial improvements in predictive performance, real-time control, and process optimization. This growing body of literature validates the approach presented in this work, which integrates IoT technologies with ensemble regression models such as Random Forest and Gradient Boosting [9, 10, 11] to forecast milk consumption patterns. The use of multiple temporal and contextual features, as shown in the present study, further enhances prediction accuracy, aligning with the best practices and quantitative gains reported in recent research.

3 Methodology

The primary objective of this project is to develop an IoT system for real-time monitoring and management of a milk dispenser. The project aims to create a sensor infrastructure capable of continuously monitoring milk levels and the internal temperature of the dispenser. To achieve this, the system integrates the following sensors: a DHT11 temperature sensor, selected for its simplicity, low cost, and adequate precision for monitoring storage conditions, and a VL53L1X time-of-flight (ToF) distance sensor, which measures the distance from the sensor to the milk surface. This non-contact optical method allows for accurate volume estimation by comparing the measured distance with the known internal height of the dispenser, avoiding the mechanical complexity and calibration issues associated with load cells.

The final objective is to implement machine learning algorithms capable of analyzing sensor data to forecast milk consumption and detect potential anomalies, such as irregular patterns that could indicate blockages or malfunctions, thereby enabling timely maintenance and uninterrupted operation.

This IoT system is designed to provide real-time tracking of milk levels and temperature, detect potential blockages, and predict consumption patterns. The architecture of this system consists of three components:

- Hardware: used for the data collection.
- Cloud storage: used for storing the data.
- Desktop application: used for user interaction.

The system architecture can be seen in Fig. 1 above, which illustrates the integration of sensors with the ESP32 microcontroller, cloud communication, and user visualization. In addition to the sensors implemented in the prototype, the diagram also includes a pH sensor and a weight sensor, which were added to reflect the potential future extension of the application and the possibility to integrating additional measurement capabilities in later development stages.

The hardware layer acts as the core processing unit and is connected to the necessary sensors to collect data: the distance sensor measures the level of milk in the dispenser and the temperature sensor monitors the temperature inside the dispenser. The sensors are linked to the microcontroller, which manages raw data and prepares it for transmission. The microcontroller processes the collected data, then transmit it to the cloud. After the data is gathered using the system, it is recorded inside Firebase platform using WI-FI module. The desktop application retrieves real-time sensor data from Firebase and allows the user to visualize the output in an intuitive interface.

3.1.Implementation

The electrical circuit presented in Fig. 2 illustrates the integration of the DHT11 temperature and the VL53L1X time-of-flight distance with the ESP32 development board.

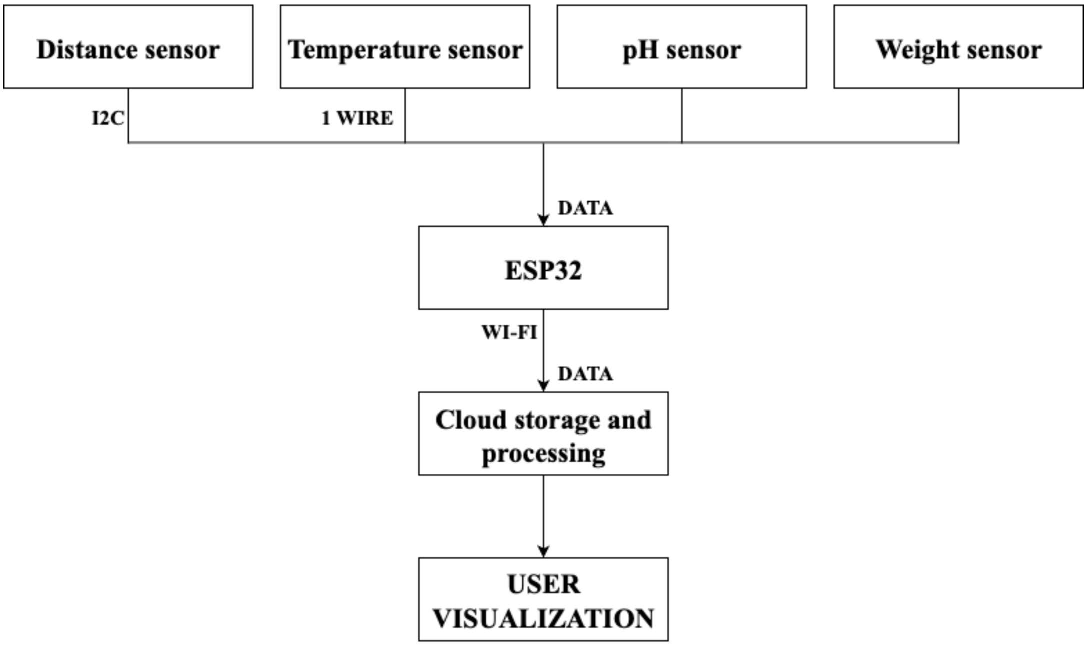

Figure 1. System architecture

The DHT11 sensor utilizes a single-wire digital communication protocol, enabling reliable temperature and humidity data acquisition with minimal wiring complexity. In contrast, the VL53L1X operates via the Inter-Integrated Circuit (I²C) communication protocol, which allows efficient data transfer through its Serial Data (SDA) and Serial Clock (SCL) lines.

The ESP32 serves as the central processing unit, providing the necessary voltage supply, ground reference, and data communication lines for both sensors. This configuration ensures accurate environmental sensing while maintaining compatibility with the microcontroller's digital and I²C interfaces.

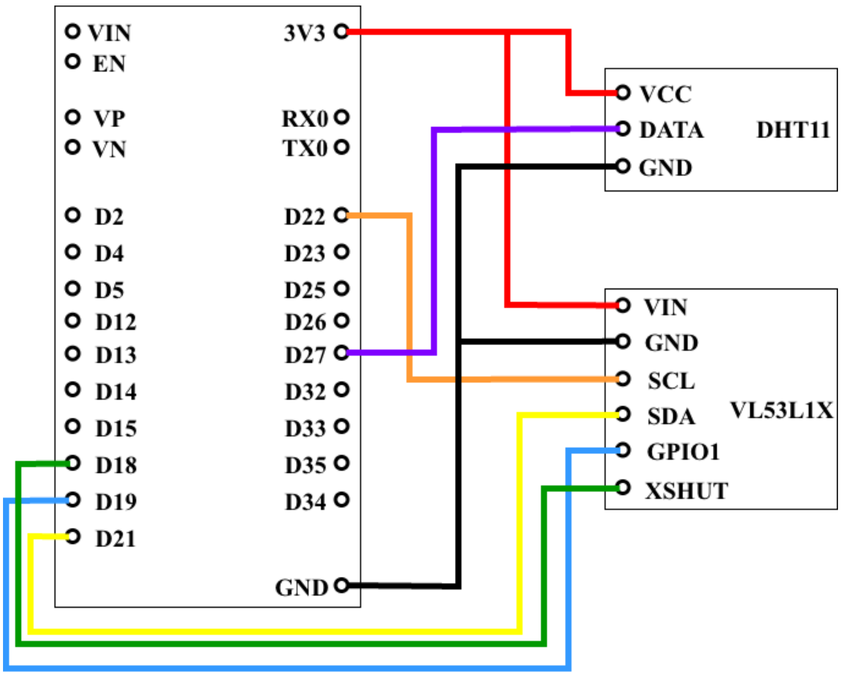

Figure 2. Electrical circuit

The real hardware implementation of the proposed system is illustrated in Fig. 3 below.

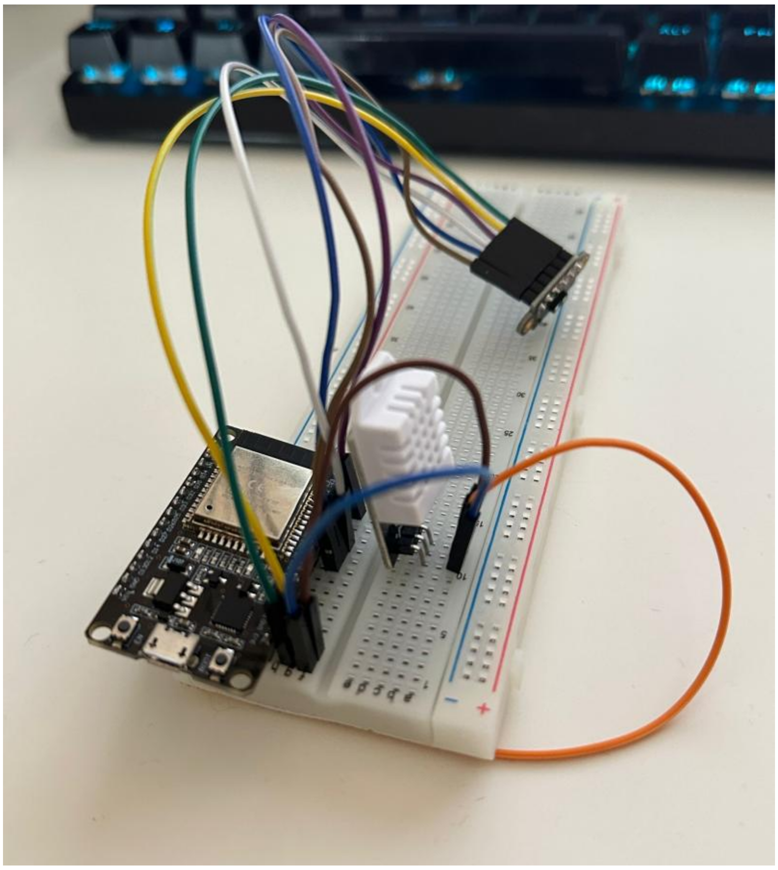

Figure 3. Real-life hardware setup with ESP32 and sensors

Fig. 4 illustrates the complete workflow for predicting milk consumption using machine learning. The process begins with data collection from Firebase, followed by preprocessing, where only events labeled “Buy” are retained and subsequently grouped by day to compute daily quantities. In cases where a date contains no data, a value of zero is assigned to preserve the continuity of the time series. Feature extraction is then performed, incorporating attributes such as the day of the week and weekend indicators. Based on these features, a regression model is selected. When Random Forest is applied, multiple trees are trained in parallel and their outputs are averaged, while Gradient Boosting trains trees sequentially, each improving upon the previous iteration. The final predictions are stored and subsequently evaluated using standard performance metrics.

Random Forest is an ensemble learning method that constructs multiple independent decision trees using the bagging technique. These trees are trained in parallel and subsequently combined, a strategy that reduces overfitting and enhances prediction accuracy, while also providing faster training. In contrast, Gradient Boosting builds trees sequentially, where each new tree corrects the errors of its predecessors. This approach is more effective in capturing complex patterns but generally incurs higher computational costs. In the context of milk consumption forecasting, both algorithms are applied to historical data in order to improve demand prediction and system reliability.

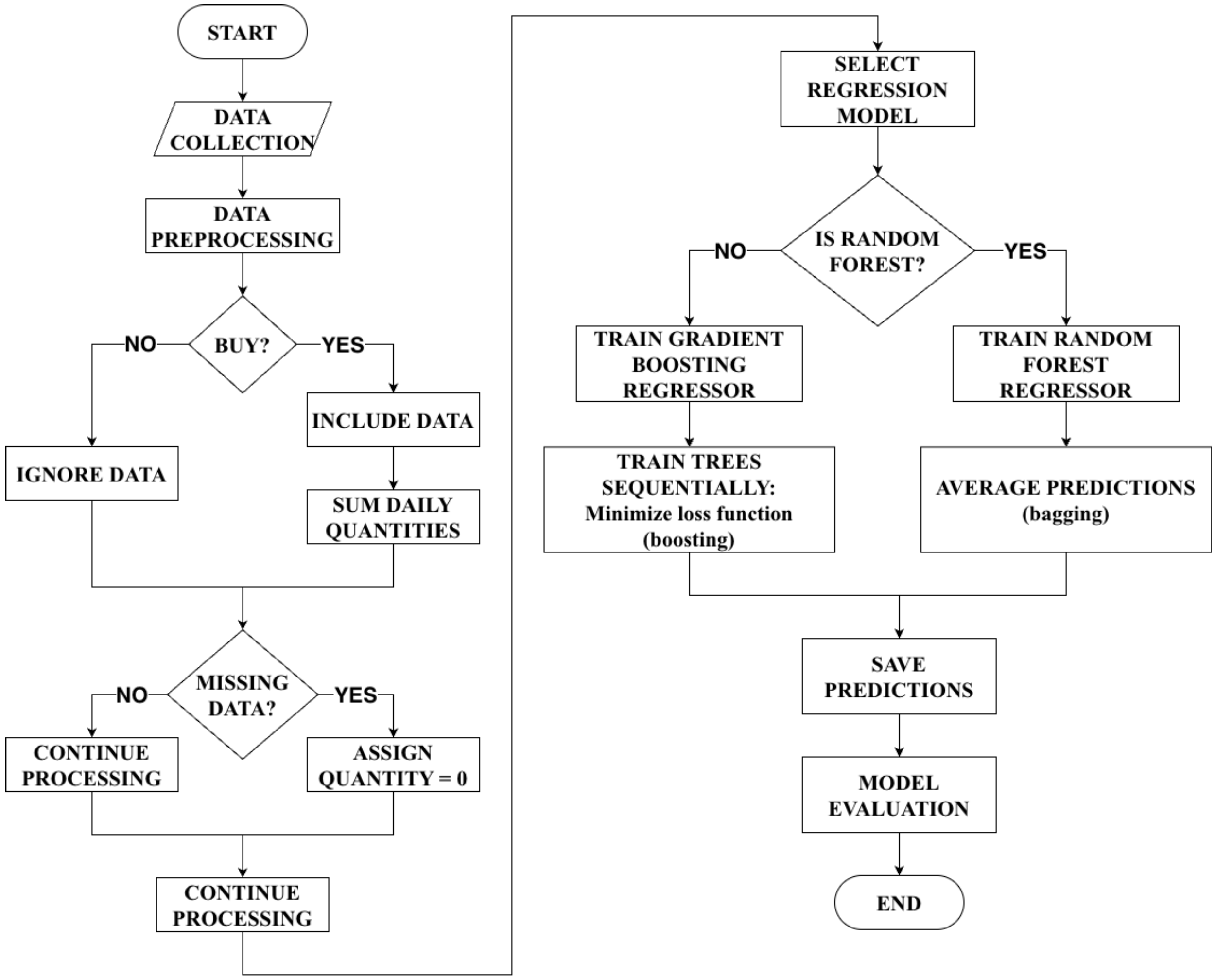

Figure 4. Regression model flow

4 Results

The prediction task was conducted using two distinct feature sets. The first set consisted of the day of the week, encoded as an integer ranging from 0 (Monday) to 6 (Sunday), and a Boolean variable indicating whether the day corresponded to a weekend or a weekday. The performance of both algorithms based on these features is illustrated in the presented results. The findings indicate that both models achieve reasonably accurate forecasts, successfully capturing the general trend of the actual values even with a minimal feature set. The discrepancies between predicted and observed values remain relatively small; however, Gradient Boosting demonstrates superior performance in most cases, exhibiting a stronger capacity for generalization.

Table 1 compares actual recorded daily milk consumption to the corresponding predictions given by the Random Forest and Gradient Boosting models over a one-week period. The table demonstrates the models' ability to estimate actual consumption trends using historical data and specified temporal variables. By comparing the genuine values to the anticipated outputs, it is able to visually analyze each algorithm's accuracy and consistency, as well as find tiny differences between observed and estimated values. This comparison provides an early assessment of the prediction performance of both models.

Table 1. Prediction - two features

Date	True value(liters)	Random Forest (liters)	Gradient Boosting (liters)
5/29/2025	75	72.15	72.14
5/30/2025	61	70.32	70.23
5/31/2025	79	76.31	76.43
6/1/2025	68	72.24	72.00
6/2/2025	75	72.94	73.00
6/3/2025	70	71.94	71.93
6/4/2025	78	70.85	70.86

The graphical representation from Fig. 5, emphasizes the close alignment between predicted and actual values across the tested days, although minor deviations are still present.

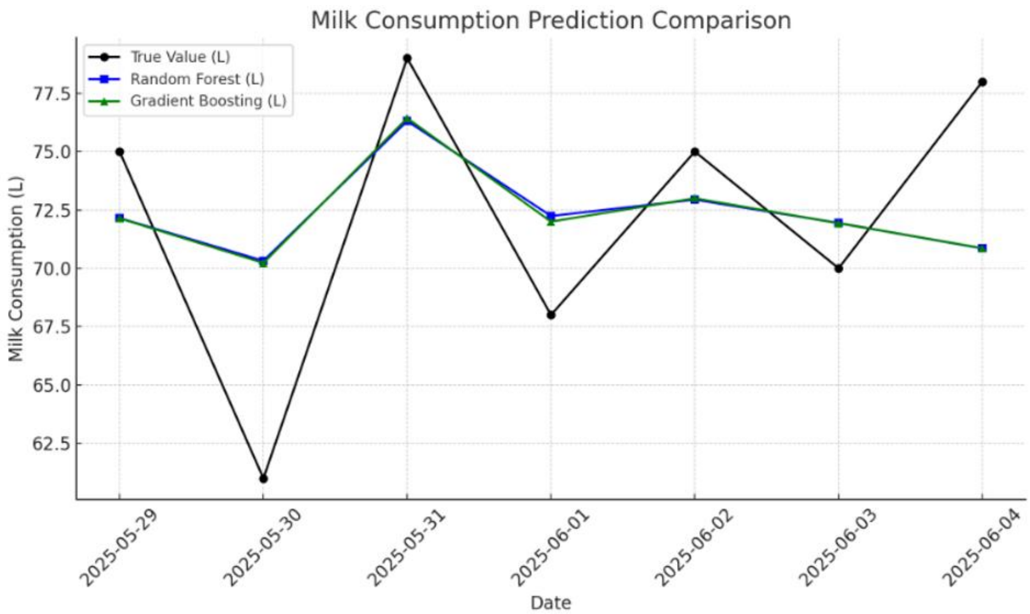

Figure 5. True vs predicted values - two features

To evaluate the performance of the models, several error metrics were computed, as summarized in Table 2.

Table 2. Metrics for two features

	MAE	RMSE	R^2
Random Forest	4.32	5.06	0.26
Gradient Boosting	4.27	5.01	0.28

The Mean Absolute Error (MAE) obtained by both algorithms is approximately 4.3, indicating relatively small deviations between predicted and actual values. The Root Mean Squared Error (RMSE) is slightly higher than the MAE, reflecting the presence of occasional larger prediction errors. Gradient Boosting achieves a marginally lower RMSE compared to Random Forest, suggesting a reduced frequency of large errors. Although the coefficient of determination (R^2) values is relatively low, they still indicate that the models are able to capture part of the variance in the data, with Gradient Boosting providing a modest improvement in explanatory power.

To enhance model performance, a third feature, day of the month, was incorporated to provide additional temporal context. This feature enables the models to better capture periodic patterns, resulting in a noticeable improvement in prediction accuracy. Gradient Boosting continues to slightly outperform Random Forest, demonstrating a stronger ability to exploit the additional information. The visual comparison highlights that including the day of the month allows both models to adapt more effectively to daily variations. Table 3 compares the actual daily milk consumption values with the predictions generated by both models using the enriched three-feature configuration.

Table 3. Prediction - three features

Date	True value (liters)	Random Forest (liters)	Gradient Boosting (liters)
5/29/2025	75	73.59	75.94
5/30/2025	61	64.58	62.09
5/31/2025	79	74.13	77.43
6/1/2025	68	66.65	68.87
6/2/2025	75	75.09	74.99
6/3/2025	70	71.97	72.61
6/4/2025	78	77.06	77.72

Including this feature results in a noticeable improvement in prediction accuracy across most cases. Furthermore, the slight performance advantage observed for Gradient Boosting indicates that this algorithm is more effective in leveraging the additional information provided by new features, thereby producing more precise predictions.

Figure 6 presents the comparison between the observed values and the predictions generated by the two algorithms when three features are included. The results clearly indicate that adding the day of the month as a feature enables the models to better capture daily variations compared to the two-feature configuration. This observation is further supported by the evaluation metrics summarized in Table 4.

Gradient Boosting consistently outperforms Random Forest across all metrics, achieving a substantially higher R^2 and significantly lower error values. These results highlight the superior capacity of Gradient Boosting to exploit the richer feature set, thereby generating more accurate and reliable predictions.

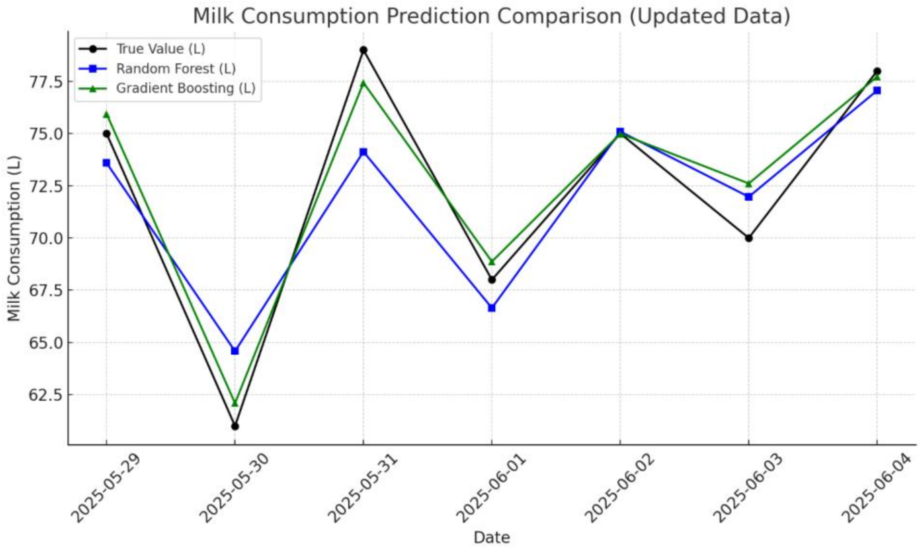

Figure 6. True vs predicted values - three features

Table 4 Metrics - three features

	MAE	RMSE	R^2
Random Forest	2.03	2.54	0.81
Gradient Boosting	1.05	1.32	0.95

To investigate the performance of the model and evaluate the relevance of the selected features, a new dataset covering the period January 2024 – June 2025 was generated. Since the prototype was not yet connected to a real milk dispenser, the training data were synthetically generated within realistic value ranges, allowing the analysis to be carried out in a controlled environment. In future work, these models will be retrained on real world operational data once the system is deployed on an actual dispenser. This extended dataset allowed the introduction of two additional features, public holiday and season, which were used to encapsulate temporal patterns that may influence milk consumption patterns. The four-feature set includes a binary variable indicating whether a given day is a public holiday, reflecting potential changes in consumer behavior on special days. The five-feature set further incorporates the season, providing a broader temporal context that may affect consumption trends. By progressively enriching the feature set, the models can better learn periodic and contextual variations, which is expected to improve predictive accuracy for both Random Forest and Gradient Boosting algorithms.

Table 5 presents the predictions obtained by the Random Forest model for a one-week period, from April 2 to April 8, 2025. The results are shown for configurations using two to five features. When only two features are used, the discrepancies between predicted and actual values are evident. For example, on April 2, the observed value was 85.99 liters, while the model predicted 72.23 liters, indicating a substantial underestimation and highlighting the limitations of a basic temporal structure. In contrast, the five-feature set produces a prediction of 85.42 liters for the same day, demonstrating a significant improvement in accuracy and a stronger ability to capture daily consumption patterns.

Table 5. Random Forest

Date	True value (liters)	Two features (liters)	Three features (liters)	Four features (liters)	Five features (liters)
4/2/2025	85.99	72.23	87.02	86.99	85.42
4/3/2025	77.17	68.18	80.13	80.35	79.73
4/4/2025	77.9	67.06	81.82	81.83	79.8
4/5/2025	81.92	77.92	76.14	76.29	80.18
4/6/2025	82.89	77.45	77.18	77.16	79.89
4/7/2025	63.76	68.34	65.2	65.23	63.69
4/8/2025	64.02	67.96	65.53	65.54	63.55

Table 6 summarizes the evaluation metrics for the Random Forest model across the different feature sets. When only two features are used, the model performs poorly, with a low R^2 value indicating its inability to capture meaningful patterns in the data. Adding a third feature leads to a substantial improvement: MAE decreases from 7.36 to 3.19, and R^2 increases from 0.02 to 0.8. Adding the holiday feature (fourth feature) yields similar performance to the three-feature configurations, which confirms the fact that holidays may not influence in a drastic way the weekly predictions. The best performance is achieved with the five-feature set, which underline the importance of including seasonal information for more accurate forecasting.

Table 6. Metrics – Random Forest

Number of features	MAE	RMSE	R^2
Two	7.36	8.19	0.02
Three	3.19	3.69	0.80
Four	3.21	3.69	0.80
Five	1.47	1.80	0.95

Table 7 presents the corresponding predictions obtained using the Gradient Boosting model over the same period and feature configurations. As with Random Forest, the use of only two features results in underperformance, particularly on April 2, where the predicted value significantly deviates from the observed one. The inclusion of a third feature improves the alignment between predicted and actual values, while the addition of holiday and seasonal features leads to further refinements in accuracy. The five-feature configuration produces the most balanced and accurate predictions, with values closely matching the true consumption data on multiple days.

Table 7. Gradient Boosting

Date	True value (liters)	Two features (liters)	Three features (liters)	Four features (liters)	Five features (liters)
4/2/2025	85.99	72.23	84.74	84.52	84.77
4/3/2025	77.17	68.22	81.52	81.66	81.55
4/4/2025	77.9	66.99	81.69	81.68	81.54
4/5/2025	81.92	77.93	76.92	77.58	77.22
4/6/2025	82.89	77.42	76.41	76.46	77.41
4/7/2025	63.76	68.37	64.86	65.0	64.88
4/8/2025	64.02	67.94	64.87	64.65	64.79

Table 8 shows the evaluation metrics for the Gradient Boosting model. Similar to Random Forest, performance with two features is weak, reflected by high error values and a low R^2. Adding the third feature leads to a substantial improvement in both MAE and R^2, while the inclusion of holiday and seasonal features further stabilizes the model's performance, yielding the best overall results with the five-feature configuration.

Table 8 Metrics – Gradient Boosting

Number of features	MAE	RMSE	R^2
Two	7.35	8.20	0.02
Three	3.26	3.85	0.78
Four	3.20	3.76	0.79
Five	3.04	3.54	0.82

In conclusion, both Random Forest and Gradient Boosting models benefit significantly from the inclusion of additional temporal and contextual features. While Random Forest achieves slightly better performance with the full feature set, Gradient Boosting demonstrates stable behavior even with the fewer features, confirming the effectiveness of ensemble learning techniques for forecasting consumption patterns in time-series data.

Figure 7 compares the evaluation metrics of both models across progressively richer feature sets. The results highlight that the inclusion of the day-of-month and seasonal features substantially reduces MAE and RMSE, while simultaneously increasing R^2. Although Gradient Boosting demonstrates stable performance with fewer features, Random Forest slightly outperforms it when the complete feature set is employed, achieving the most accurate predictions overall.

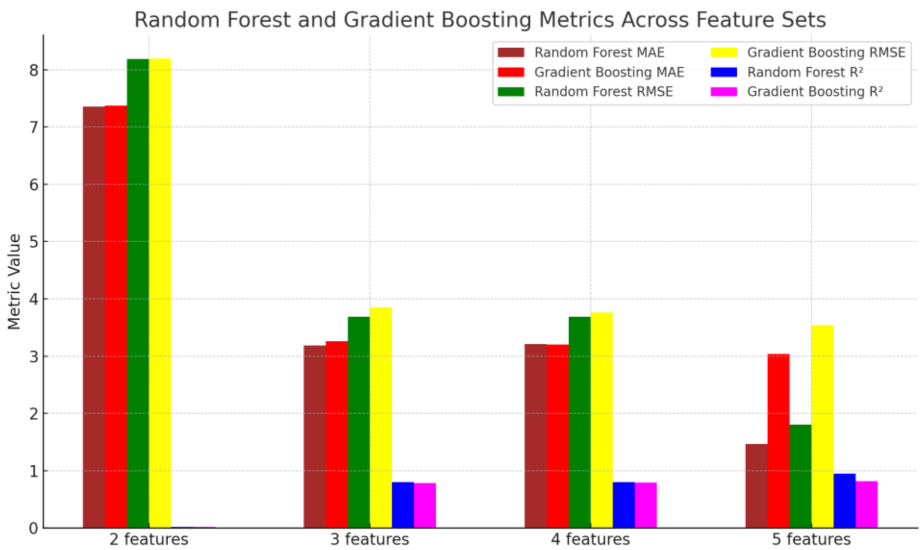

Figure 7. Performance metrics comparison

4.1. Prototype Implementation

To validate and test the functionality of the implemented system, a basic prototype was built using a container filled with water and the sensors mounted above it. The purpose was to simulate variations in liquid volume and confirm that the system accurately detects and displays the quantity in the monitoring application.

The prototype includes the following elements and could be seen in the Fig. 8 below:

- A container partially filled with water used to simulate the milk tank.
- VL53L1X sensor placed above the container connected to the microcontroller.

It is important to note that although the system architecture includes both a temperature and a distance sensor, the prototype used in this testing stage implements only the distance sensor. The temperature sensor was planned to be tested on an actual milk dispenser, but this step could not be implemented in time; therefore, the displayed temperature value represents an approximative reference corresponding to typical dispenser conditions.

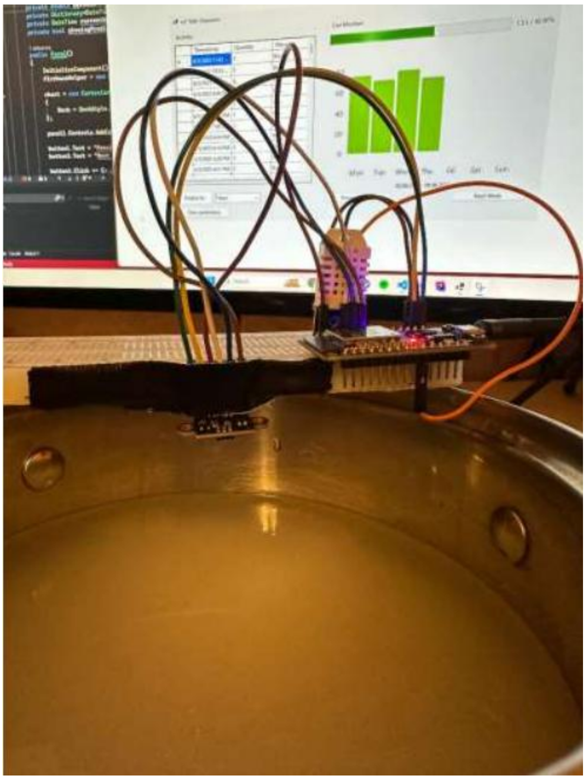

Figure 8. Initial setup

To observe the changes in quantity, measurements were taken at two stages: before and after consumption. Initially, the container was filled with 1.5 liters of water. The sensor measured the distances from its position to the water surface. By using the dimensions of the container, the measured distance was converted into an estimated volume of liquid.

The images below illustrate the experimental technique used to validate the system's ability to detect and display liquid level changes in real time. Figure 9 shows the initial state of the container before any modifications were made. To simulate real usage conditions, 0.5 liters of water were manually removed from the container. After this change, the sensor immediately detected the new distance between the sensor and the liquid surface, and the application updated the displayed value accordingly, as illustrated in Figure 10. This experiment confirmed that the system can accurately register and reflect even small variations in liquid level, demonstrating both the precision of the sensor and the responsiveness of the data transmission and visualization process.

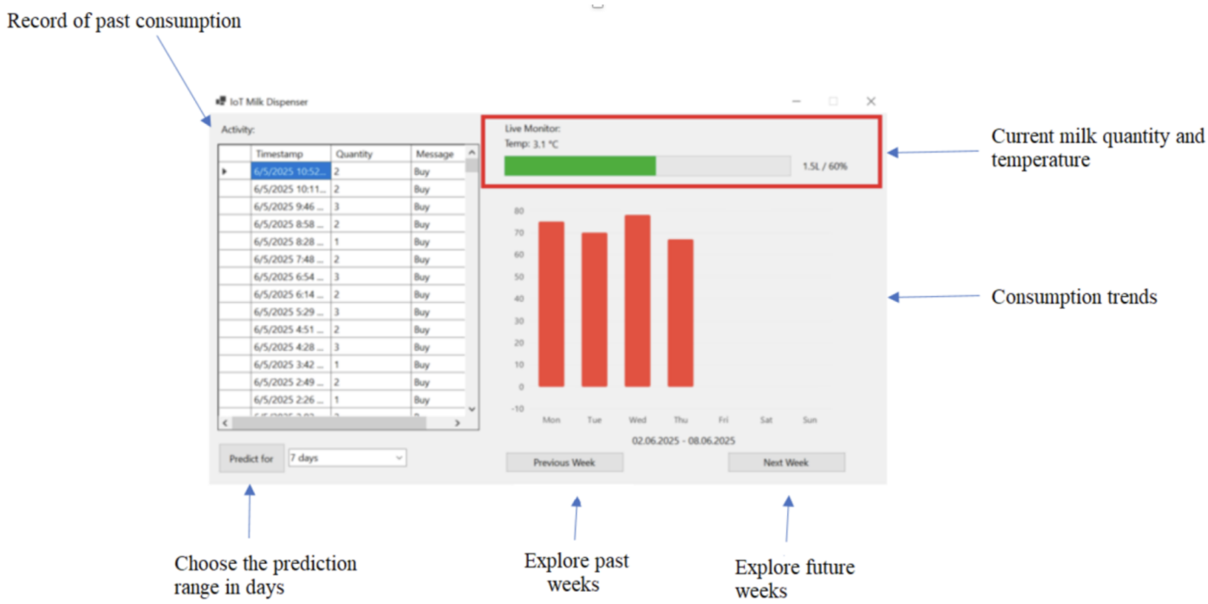

Figure 9. Initial state of the container

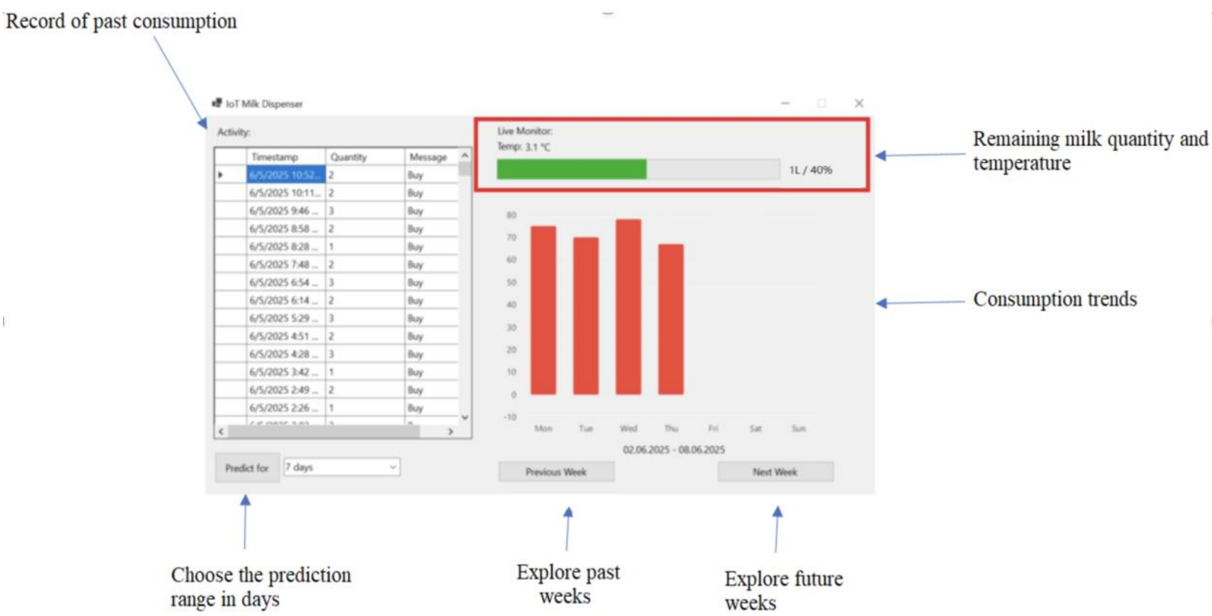

Figure 10. Final state of the container

5 Conclusions

This article explored the design and prototyping of an IoT-based system for monitoring the level of milk in a dispenser. The developed system successfully integrated multiple sensors with an ESP32 microcontroller, Firebase for cloud-based synchronization, and a dedicated application for real-time visualization. The experimental evaluation of the prediction models demonstrated that the inclusion of more than three features significantly improved the accuracy of the forecasts. Both Random Forest and Gradient Boosting benefited from the richer feature set, but Gradient Boosting consistently outperformed Random Forest, achieving higher R^2 values and lower error metrics. These results confirm that the integration of additional temporal and contextual information enhances the capacity of the models to anticipate milk consumption patterns with higher reliability.

Acknowledgment

This research was conducted as part of the collaborative work of three authors within the Embedded Systems master's program at the Faculty of Engineering, "Lucian Blaga" University of Sibiu, Romania. The authors gratefully acknowledge the support and guidance provided throughout the development of this study.

References

[1] Dataintelo. (2024). Milk Vending Machines Market – Global Industry Analysis, Size, Share, Growth, Trends, and Forecast 2024–2032. Dataintelo

[2] Rabiya Abbasi, Pablo Martinez, Rafiq Ahmad, The digitization of agricultural industry – a systematic literature review on agriculture 4.0, Smart Agricultural Technology, Volume 2, 2022, 100042, ISSN 2772-3755, https://doi.org/10.1016/j.atech.2022.100042.

[3] M. Ayaz, M. Ammad-Uddin, Z. Sharif, A. Mansour and E. -H. M. Aggoune, "Internet-of-Things (IoT)-Based Smart Agriculture: Toward Making the Fields Talk," in IEEE Access, vol. 7, pp. 129551-129583, 2019, doi: 10.1109/ACCESS.2019.2932609.

[4] Haozhe Zhang, Long He, Francesco Di Gioia, Daeun Choi, Antonio Elia, Paul Heinemann, LoRaWAN based internet of things (IoT) system for precision irrigation in plasticulture fresh-market tomato, Smart Agricultural Technology, Volume 2, 2022, 100053, ISSN 2772-3755, https://doi.org/10.1016/j.atech.2022.100053.

[5] Kaneez Fizza, Abhik Banerjee, Dimitrios Georgakopoulos, Prem Prakash Jayaraman, Ali Yavari, nas Dawod, An Inexpensive AI-Powered IoT Sensor for Continuous Farm-to-Factory Milk Quality Monitoring, Sensors 2025, 25, 4439, https://doi.org/10.3390/s25144439.

[6] V. Manibabu and M. Gomathy, "Predictive Modeling of Milk Production Using Artificial Intelligence and Machine Learning Techniques," International Journal of Computer Techniques, vol. 12, no. 3, May-June 2025.

[7] Syrmos, E. et al. (2023). An IoT Monitoring System for Dairy Products. In: Rabie, K., Kharel, R., Mohjazi, L., Elganimi, T., Khan, W.U. (eds) IoT as a Service. IoTaaS 2022. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 506. Springer, Cham. https://doi.org/10.1007/978-3-031-37139-4_2.

[8] R. D. Prasad and S. Chidambaranathan, "Digitizing Milk Collection Stage with IoT-Enabled System for Quality Assessment using ML," Journal of Information Systems Engineering and Management, vol. 10, no. 29s, 2025, https://doi.org/10.52783/jisem.v10i29s.4604.

[9] Henzel, J., & Sikora, M. (2020). Gradient Boosting Application in Forecasting of Performance Indicators Values for Measuring the Efficiency of Promotions in FMCG Retail. Silesian University of Technology, Faculty of Automatic Control, Electronics and Computer Science.

[10] Ayaru, L., Ypsilantis, P.-P., Nanapragasam, A., Choi, R. C.-H., Thillanathan, A., Min-Ho, L., & Montana, G. (2021). Prediction of Outcome in Acute Lower Gastrointestinal Bleeding Using Gradient Boosting.
[11] Boldini, D., Grisoni, F., Kuhn, D., Friedrich, L., & Sieber, S. A. (2023). Practical guidelines for the use of gradient boosting for molecular property prediction. Journal of Cheminformatics, 15(1), Article 35.