Jurnal Tekno Insentif | Vol. 19 | No. 2 | Desember 2025
DOI : https://doi.org/10.36787/jti.v19i2.2087 Halaman 223-238

Multi-Horizon Prediction Of Broiler Mortality With Decision Tree And SVM: A Case Study In Small-To-Medium Farms In Sukabumi

BUDI JEJEN ZAENAL ABIDIN[1], BAMBANG SOMANTRI WIJAYA[2], GHAIDA SANDIE HAYATUS SAHLA[1]

[1]Program Studi Sistem Informasi Universitas Sains Indonesia, Indonesia
[2]Program Studi Manajemen Universitas Sains Indonesia, Indonesia
Email: budi.jejen@lecturer.sains.ac.id

ABSTRAK

Penelitian ini mengembangkan model machine learning untuk memprediksi mortalitas harian (jumlah kematian) ayam broiler 1–7 hari ke depan secara multi-horizon (H+1–H+7) menggunakan data 12 kandang di Sukabumi selama lima siklus produksi (Juli 2024–Juli 2025). Data diprapróses melalui imputasi nilai hilang (Random Forest), penanganan outlier (IQR dan winsorizing), normalisasi Z-score, serta seleksi fitur (Pearson dan ReliefF). Support Vector Regression (SVR) dan Decision Tree Regression (DTR) dibandingkan dengan hasil menunjukkan SVR unggul untuk prediksi jangka pendek (H+1–H+2; R^2 = 0,842–0,760), tetapi performanya menurun pada horizon yang lebih panjang. Sebaliknya, DTR lebih stabil pada horizon menengah–panjang (H+5–H+7; $R^2 \approx$ 0,683–0,696). Faktor dominan yang berkaitan dengan mortalitas adalah umur dan bobot rata-rata, serta kondisi kandang seperti ventilasi/kecepatan angin, kepadatan tebar, NH_3, dan suhu. Evaluasi dilakukan dengan repeated holdout 70/30 (10 repetisi) dan 5-fold cross-validation pada data latih, mendukung prototipe sebagai peringatan dini.

Kata kunci: ayam broiler, *decision tree regression*, mortalitas, *multi-horizon forecasting, support vector regression*

ABSTRACT

This study develops a machine-learning model to predict daily broiler mortality (death counts) 1–7 days ahead using a multi-horizon approach (H+1–H+7), based on data from 12 broiler houses in Sukabumi across five production cycles (July 2024–July 2025). Data were preprocessed using missing-value imputation (Random Forest), outlier handling (IQR and winsorizing), Z-score normalization, and feature selection (Pearson correlation and ReliefF). Support Vector Regression (SVR) and Decision Tree Regression (DTR) were compared. Results show that SVR outperformed DTR for short-term prediction (H+1–H+2; R^2 = 0.843–0.760), but its performance declined at longer horizons. In contrast, DTR was more stable for medium-to-long horizons (H+5–H+7; $R^2 \approx$ 0.683–0.696). Dominant factors associated with mortality included age and average body weight, as well as housing conditions such as ventilation/wind speed, stocking density, NH_3, and temperature. Evaluation used repeated 70/30 holdout (10 repetitions) and 5-fold cross-validation on the training data, supporting a prototype as an early warning tool.

***Keywords**: broiler chicken, decision tree regression, mortality, multi-horizon forecasting, support vector regression*

1. INTRODUCTION

Broiler chicken raising is a crucial industry in supplying economical animal protein. Broiler mortality continues to be a substantial worldwide challenge, resulting in economic losses, diminished production, and issues over animal welfare. Research has shown that factors such as stocking density **(Zabir et al., 2021)**, early rearing circumstances (Yerpes et al., 2020), heat stress **(Liu et al., 2020)**, and weak biosecurity practices **(Riber & Wurtz, 2024; Ziebe et al., 2025)** influence mortality rates. Recent studies indicate that adherence to biosecurity measures directly correlates with production performance and mortality rates, with farms exhibiting inadequate biosecurity experiencing elevated death rates and increased antibiotic use.

Precision livestock farming is advancing, using technology such as sensors, big data, and machine learning (ML) to monitor animal health and forecast mortality risks. Numerous research studies have used machine learning for poultry disease identification **(Kader et al., 2021),** broiler development modeling **(Ahmad, 2009)**, and the integration of big data for contemporary farm management **(Neethirajan, 2020)**. Moreover, current evaluations indicate that deep learning is emerging as a prominent trend in enhancing the precision of poultry health forecasts **(Shwetha et al., 2024)**. In small- to medium-sized farms in Indonesia, traditional monitoring techniques prevail, mainly depending on subjective experience and wasteful manual inspections (Adelia & others, 2024). This issue complicates the early identification of beginning symptoms.

To surmount these constraints, machine learning methodologies are being progressively embraced. These methods may augment efficiency, diminish expenses, and elevate animal welfare **(Neethirajan, 2020)**. Regression methods, like Decision Tree Regression (DTR) and Support Vector Regression (SVR), are especially appropriate given that the goal variable is the count of chicken fatalities. Despite the use of machine learning in the livestock industry, research on multi-horizon forecasting (1–7 days ahead) for broiler mortality is limited, especially for small- to medium-scale farms in Indonesia.

Previous works have applied ML in poultry farming, such as disease detection **(Kader et al., 2021; Kholil et al., 2022)**, growth prediction using regression **(Ahmad, 2009)**, mortality prediction with Naïve Bayes and C4.5 **(Imam Baihaqi et al., 2019)**, and temperature control using fuzzy systems **(Darmawi et al., 2020)**. However, few studies directly compare the performance of decision tree regression and support vector regression (SVR) in predicting broiler mortality, especially with a multi-horizon prediction approach (1–7 days ahead).

Based on these problems, this study was conducted as a case study in Sukabumi Regency to answer three main questions: (1) How can mortality prediction models be built using DTR and SVR that achieve strong evaluation performance (MAE, RMSE, R^2, CV-RMSE)? (2) What factors most significantly influence broiler mortality? (3) How can model validation be conducted to ensure accuracy in real-world scenarios?

The aims of this work are to (i) create a dependable machine learning-based mortality prediction model, (ii) ascertain the primary determinants affecting broiler mortality, and (iii) construct a prototype prediction system as a preliminary measure for field application. This research primarily contributes (1) the implementation of multi-horizon forecasting (H+1 to H+7) in small- to medium-scale poultry farms, (2) a systematic comparison of DTR and SVR to identify the most effective algorithm for mortality prediction, and (3) the creation of a MATLAB-based graphical user interface (GUI) prototype as a proof of concept for an early warning system. This research enhances the methodology of data mining by examining multi-horizon performance, dominant feature analysis, and system prototyping, while also offering practical solutions for the application of smart poultry farming in Indonesia.

2. METHODS

2.1 Research Stages

This research used the CRISP-DM (Cross-Industry Standard Process for Data Mining) paradigm, including six stages: (1) Business Understanding, (2) Data Understanding, (3) Data Preparation, (4) Modeling, (5) Evaluation, and (6) Deployment. This framework was chosen because of its systematic approach, which includes defining requirements and implementing predictive models. Figure 1 depicts the comprehensive study process, highlighting key variables such as chicken age, daily body weight, daily increase, stocking density, temperature, wind speed, NH_3 concentration, and daily mortality, as established by prior studies. The prediction aim was established as daily mortality, spanning from H+1 (next-day mortality) to H+7 (seven days in advance).

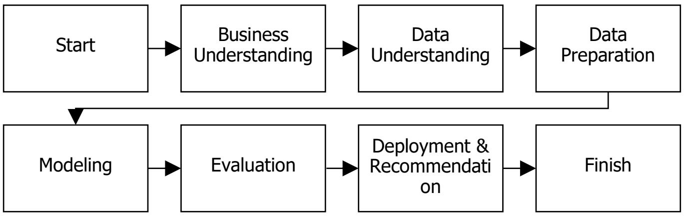

Figure 1. Research Stages

2.2 Dataset

The information was obtained from 12 broiler farms in Sukabumi Regency throughout five raising seasons (July 2024-July 2025). It comprises two categories of predictor variables: production variables (age, average weight, and daily feed intake) and environmental variables (temperature, humidity, wind speed, ammonia concentration (NH_3), and stocking density). To enable multi-horizon forecasting, the original daily mortality series was transformed into seven horizon-specific targets. For each observation at day t, the target variables were defined as mortality at day (t+h), where $h \in \{1, 2, 3, 4, 5, 6, 7\}$.

Table 1. Characteristics of Farm Data

Attribute	Description and Attribute Type
Production Characteristics	
Chicken Age	Day count since chickens were first reared (numeric)
Average Daily Weight	Average chicken body weight per day (grams) (numeric)
Daily Gain (Weight Increase)	Difference in body weight per day (daily gain) (numeric)
Daily Mortality	Number of chickens that died per day (numeric)
Environmental Characteristics	
Environmental Temperature	Temperature inside the poultry house (°C) (numeric)
Wind Speed	Airflow/air speed inside the poultry house (m/s) (numeric)
Humidity	Relative humidity inside the poultry house (%RH) (numeric)
NH_3	Ammonia concentration (ppm)(numeric)
Stocking Density	Number of chickens per square meter in the poultry house (numeric)

House Identity	Poultry house identifier (nominal)

This yielded seven supervised learning datasets (H+1 to H+7). Observations at the end of each production cycle that did not have valid future mortality values for a given horizon were excluded to avoid label leakage across horizons. The same set of predictor variables was used for all horizons to ensure comparability of model performance across lead times.

Table 2. List of Input Attributes Used in This Study and References from Previous Research

Attribute	Previous Studies Using This Attribute
Stocking Density	**(Zabir et al., 2021)**
Chicken Age	**(Yerpes et al., 2020)**
Average Daily Weight	**(Ahmad, 2009; Riber & Wurtz, 2024)**
Daily Gain (Weight Increase)	**(Ahmad, 2009; Riber & Wurtz, 2024; Zabir et al., 2021)**
Environmental Temperature	**(Darmawi et al., 2020; Kim et al., 2021; Liu et al., 2020; Pirompud et al., 2024)**
Wind Speed	**(Alves et al., 2024; Elghardouf et al., 2023)**
NH_3 (Ammonia)	**(Cruz et al., 2024; Jainonthee et al., 2025; Shwetha et al., 2024; Wei et al., 2015)**
Humidity	**(Harrison et al., 1958; Kim et al., 2021; Wei et al., 2015)**

Table 1 summarizes the dataset features, including essential information on the number of data recordings, collecting times, and the observed production and environmental factors. Table 2 consolidates references from prior studies that used analogous traits, providing the conceptual foundation for variable selection in this study.

2.3 Data Preprocessing

The preprocessing stage ensured that the dataset was clean, consistent, and ready for modeling. The following steps were applied:

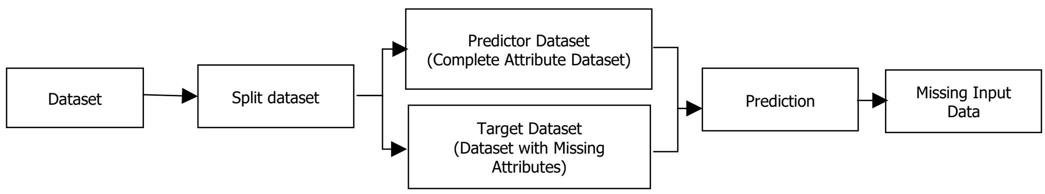

Figure 2. Prediction-based Imputation

a. Missing data imputation was carried out using the Random Forest Imputation algorithm, which estimates missing values based on inter-variable patterns (see Figure 2).
b. Outlier detection and handling were performed using the Interquartile Range (IQR) method and winsorizing, reducing extreme values without discarding important information.
c. Data normalization was applied specifically to SVR models, as this algorithm is sensitive to variable scales. Z-score normalization (Equation 1) was used, mapping each variable into a distribution with zero mean and unit variance. This ensured

comparable scales across all variables. Decision tree regression did not require normalization since it relies on attribute-based splits rather than distance metrics.

$$z = \frac{x - \mu}{\sigma} \tag{1}$$

d. Feature selection employed two approaches:
   - Pearson correlation (Equation 2): to measure the linear relationship between input variables (e.g., temperature, humidity, daily weight) and the mortality target. Correlation coefficients (r) range from –1 to 1, with values closer to ±1 indicating stronger relationships.

$$r_{XY} = \frac{\Sigma_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\Sigma_{i=1}^{n}(X_i - \bar{X})^2}.\sqrt{\Sigma_{i=1}^{n}(Y_i - \bar{Y})^2}} \tag{2}$$

   - ReliefF algorithm (Equation 3): to evaluate the relative importance of each attribute. ReliefF compares feature values against near-hit (same class samples) and near-miss (different class samples). Features that better distinguish these samples are assigned higher weights.

$$\mathrm{W[A]} = \mathrm{W[A]} - \frac{1}{m}\sum_{i=1}^{m} diff\big(A, X_i, NearHit(X_i)\big) + \frac{1}{m}\sum_{i=1}^{m} diff(A, X_i, NearMiss(X_i)) \tag{3}$$

These preprocessing steps ensured that the dataset used in modeling was representative, stable, and free of inconsistencies.

2.4 Modeling Algorithms

Two supervised regression models were evaluated for multi-horizon mortality forecasting: Decision Tree Regression (DTR) and Support Vector Regression (SVR). For each prediction horizon (H+1 to H+7), separate models were trained and tested using a repeated holdout protocol (70% training and 30% testing; 10 repetitions). Within each repetition, model stability was assessed by 5-fold cross-validation on the training set, and final performance was reported as the average across the 10 repetitions.

Decision Tree Regression (DTR). DTR was implemented as a CART regression tree that learns a set of hierarchical decision rules by recursively splitting the predictor space. At each node, the optimal split (feature and threshold) is selected to maximize the reduction in squared error (variance) of the response, thereby minimizing within-node mean squared error. The prediction for a new observation is obtained by routing it to a terminal leaf and returning the mean mortality value of training samples in that leaf. In this study, DTR was fitted using MATLAB fitrtree with MinLeafSize = 5 to control tree complexity and Surrogate = on to improve robustness when some predictors are missing.

Support Vector Regression (SVR). SVR was used to capture potentially nonlinear relationships between physiological/environmental predictors and daily mortality. Conceptually, SVR seeks a function $f(x)$ that is as flat as possible while allowing prediction errors within an ε-insensitive margin; errors larger than ε are penalized via slack variables, with the trade-off controlled by parameter C. Nonlinearity is handled through a kernel function, yielding a decision function of the form:

$$f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(x_i, x) + b \tag{4}$$

An RBF (Gaussian) kernel was employed:

$$K(x_i, x) = \exp\left(-\gamma \|x_i - x\|^2\right) \tag{5}$$

SVR was implemented using MATLAB fitrsvm with KernelFunction = 'rbf'. Predictors were standardized using Z-score normalization based on training statistics, and the same scaling was applied to the test set. Hyperparameters were retained at MATLAB default settings (BoxConstraint, Epsilon, and KernelScale).

2.5 Model Evaluation

Data were split into 70% training and 30% testing using a repeated holdout procedure (10 repetitions). Within each repetition, model stability was assessed using 5-fold cross-validation on the training set, and the cross-validated error was reported as CV-RMSE. Test-set performance was evaluated using MAE, RMSE, and R^2, and the reported results represent the mean performance across the 10 repetitions.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{6}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{7}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y}_i)^2} \tag{8}$$

$$CV - RMSE = \sqrt{kfoldLoss} \tag{9}$$

where y_i is the observed mortality, $\hat{y}_i$ is the predicted mortality, $\bar{y}$ is the mean of observed mortality, and n is the number of observations.

2.6 Deployment

The concluding part of CRISP-DM included the implementation of the model in an operational system. This research used a MATLAB-based graphical user interface (GUI) prototype for deployment. The system allows farmers to enter daily data (e.g., age, average weight, temperature, humidity) and obtain death predictions for timeframes H+1 to H+7. The GUI presents graphs that compare actual and predicted values, facilitating the interpretation of findings.

3. RESULTS AND DISCUSSION

3.1 Data Preprocessing Results

The preprocessing stage produced a clean and ready-to-use dataset for modeling. Missing values were successfully imputed using the Random Forest algorithm without altering the overall distribution, resulting in 2,306 valid records. As shown in Figure 3, the dataset became free of missing values after imputation.

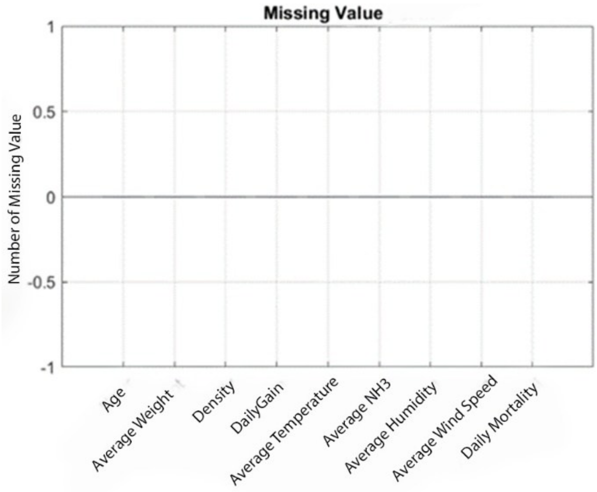

Figure 3. Visualization of Missing Values Before and After Imputation Using Random Forest

Outlier detection identified extreme values in variables such as daily gain, NH_3, humidity, and mortality. After applying winsorizing, these extreme values were reduced, making the data distribution more balanced, stable, and representative for further analysis (Figure 4).

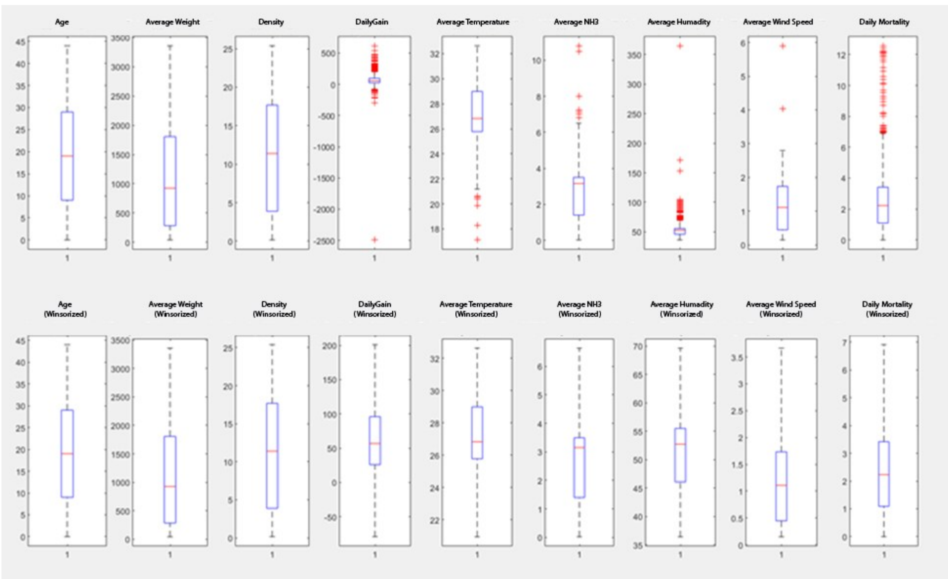

Figure 4. Results of Outlier Detection and Han-dling Using the IQR Method and Winsorizing

Feature selection was then conducted to identify the most relevant predictors of broiler mortality. Pearson correlation analysis (Table 3) revealed that age (0.86091) and average weight (0.84003) had the strongest positive correlations with mortality. Environmental variables, including average wind speed (0.77595), stocking density (0.61065), and average NH_3 concentration (0.60312), demonstrated substantial contributions, affirming that housing circumstances significantly affect chicken health. Simultaneously, average humidity (0.48827) and daily gain (0.37658) demonstrated lesser, nevertheless pertinent connections. Notably, the average temperature exhibited a negative association (–0.65189), indicating that elevated temperatures were associated with decreased mortality, necessitating additional analysis within the framework of farm management.

Table 3. Results of Pearson Correlation Between Input Variables and Broiler Chicken Mortality

Feature	Pearson Correlation
Age	0.86091
Average Weight	0.84003
Average Wind Speed	0.77595
Stocking Density	0.61065
Average NH_3	0.60312
Average Humidity	0.48827
Daily Gain	0.37658
Average Temperature	-0.65189

Table 4. Results of Decision Tree Regression Testing (Average of Repeated Holdout)

Horizon	Target	MAE	RMSE	R^2	CV RMSE
1	mortality_h1	0.202184547	0.720443723	0.813441949	0.746622929
2	mortality_h2	0.275785227	0.837225781	0.749726943	0.891543741
3	mortality_h3	0.308687993	0.855720759	0.744257505	1.006480746
4	mortality_h4	0.342087401	0.903257866	0.721288852	0.999629999
5	mortality_h5	0.381856252	0.971281085	0.683117424	1.001343641
6	mortality_h6	0.4035004	0.995401695	0.676474075	1.043266784
7	mortality_h7	0.404353999	0.977052189	0.696627015	1.083182151

Z-score normalization was applied only to the SVR model to standardize variable scales, as this algorithm is sensitive to data scaling. For decision tree regression, normalization was not necessary since the algorithm relies on threshold-based splits rather than distance metrics.

3.2 Decision Tree Regression Results

Decision tree regression was used to predict death for a duration of seven days (H+1 to H+7). The predictor variables included physiological and environmental characteristics, including age, body weight, stocking density, daily growth, temperature, ammonia levels, humidity, wind speed, and death from the previous day. Validation used a 70/30 holdout split with 5-fold cross-validation, done 10 times to guarantee robustness. The model was trained individually for each prediction horizon to account for fluctuations in predicting inaccuracy across time intervals.

The optimal performance occurred at H+1, with MAE = 0.20218, RMSE = 0.72044, and R^2 = 0.81344, indicating that the model accounted for almost 81% of the variation in actual mortality one day in advance (Table 4). Nonetheless, as the horizon broadened, performance progressively deteriorated. At H+7, R^2 decreased to 0.69663, while both MAE and RMSE increased.

Error patterns were consistent with typical multistep forecasting challenges. MAE rose from 0.27579 at H+2 to over 0.40 at H+6 and H+7. Despite this, CV-RMSE values (0.74–1.08) indicated that the model remained relatively stable across folds.

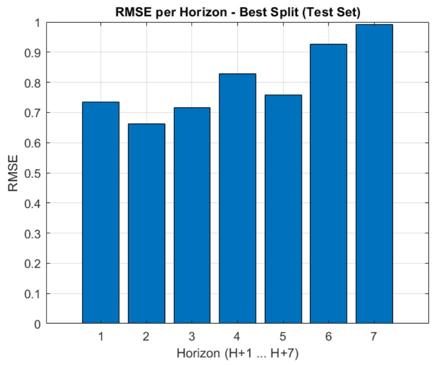

Figure 5. RMSE per Horizon – Best Split Decision Tree

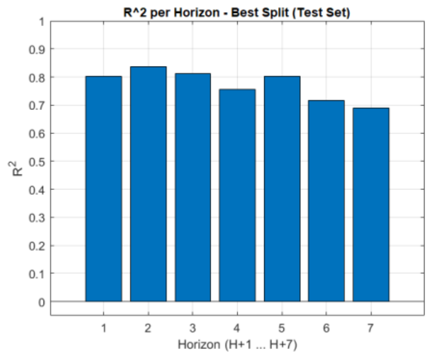

Figure 6. R2 per Horizon – Best Split Decision Tree

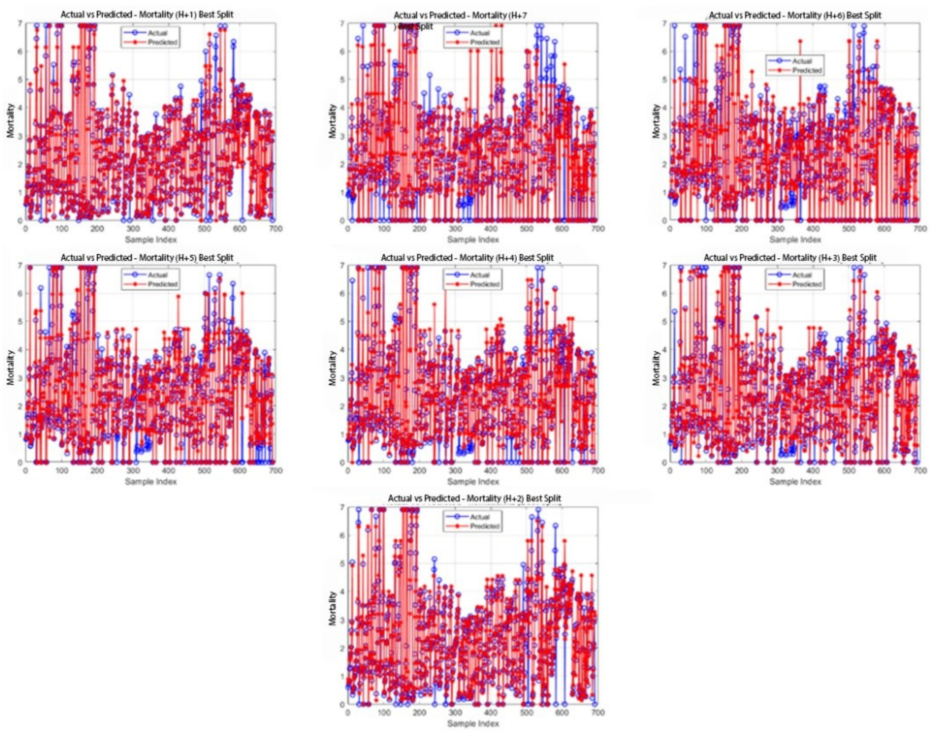

Figure 7. Actual vs. Predicted Mortality Using Decision Tree Regression

Visualizations supported these findings:

- Figure 5 showed an increasing RMSE trend from H+1 to H+7.
- Figure 6 showed decreasing R² values across horizons.

- Figure 7 (Actual vs. Predicted plots) demonstrated that predictions at H+1 closely aligned with actual values, but extended horizons exhibited more discrepancies.

Feature significance analysis (Figure 8) indicated that physiological parameters (age, body weight) and environmental factors (temperature, humidity, NH_3) were the primary predictors. Wind speed and stocking density had little impact, consistent with existing studies.

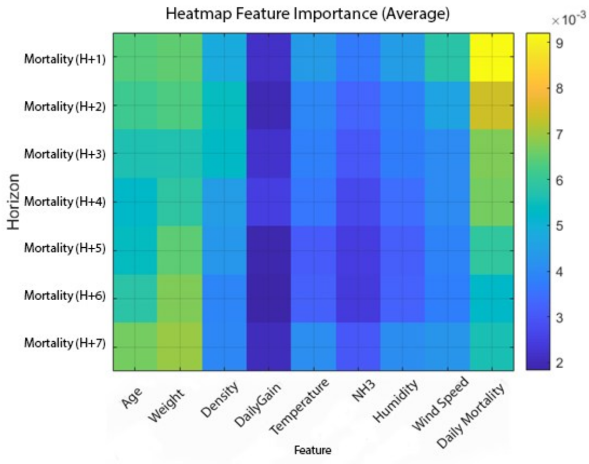

Figure 8. Feature Importance Result

3.3. Support Vector Regression Modeling

Support Vector Regression (SVR) was also applied for multi-horizon forecasting (H+1 to H+7). Like DTR, it used physiological and environmental predictors. Validation followed the same approach: 70/30 holdout, 5-fold cross-validation, and repeated training.

Table 5. Results of Support Vector Regression (SVR) Testing (Average of Repeated Holdout)

Horizon	Target	MAE	RMSE	R^2	CV RMSE
1	mortality_h1	0.264182447	0.649450066	0.84270933	0.430738832
2	mortality_h2	0.342369574	0.810334955	0.759822689	0.513298026
3	mortality_h3	0.406645691	0.89776673	0.708411661	0.541337397
4	mortality_h4	0.444169779	0.904462399	0.710237382	0.565668955
5	mortality_h5	0.493510972	0.955941589	0.6842712	0.599997915
6	mortality_h6	0.50662107	0.974221209	0.679863458	0.594383386
7	mortality_h7	0.509588916	0.939875071	0.706335391	0.591936192

SVR achieved its best performance at H+1, with MAE = 0.26418, RMSE = 0.64945, and R^2 = 0.84271, indicating it explained over 84% of the variance (Table 5). This slightly outperformed DTR in short-term prediction. However, performance dropped as the horizon increased: R^2 declined to 0.7063 at H+7, while MAE and RMSE rose above 0.50 and 0.93, respectively.

Error trends were similar to DTR:

- MAE increased steadily from H+2 onward.
- RMSE rose from 0.64945 at H+1 to nearly 0.97 at H+7.
- CV-RMSE ranged from 0.43 to 0.59, indicating that SVR was relatively stable across folds.

Visual results confirmed these patterns:

- Figure 9 showed rising RMSE per horizon.

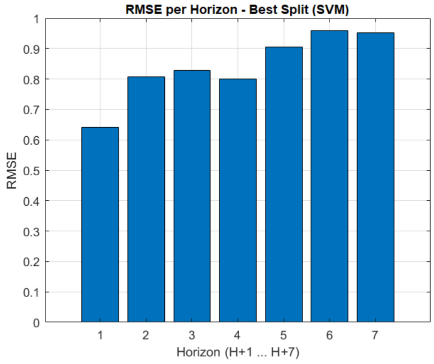

Figure 9. RMSE per Horizon – Best Split SVR

- Figure 10 displayed decreasing R^2 from ~0.84 (H+1) to ~0.70 (H+7).

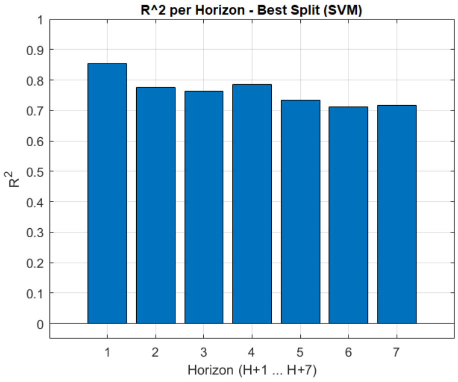

Figure 10. R2 per Horizon – Best Split SVR

- Figure 11 (Actual vs. Predicted) showed that SVR tracked actual data well at H+1 but diverged significantly at H+6 and H+7.

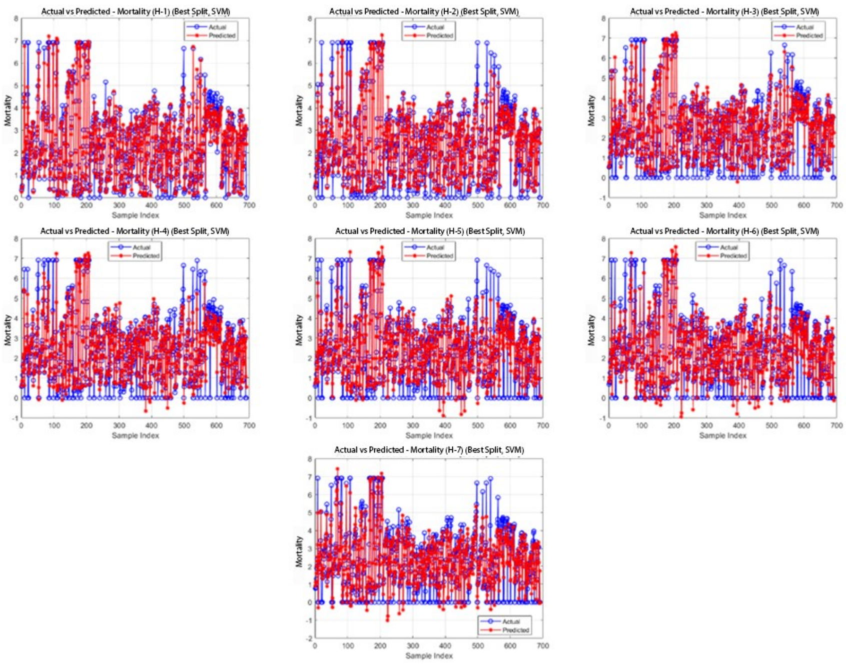

Figure 11. Plot Actual vs. Predicted

3.4 Deployment and Recommendation

The deployment stage in this study was realized through the development of a simple prototype in the form of a graphical user interface (GUI) built in MATLAB. This prototype was designed to display daily mortality prediction results based on relevant input variables such as chicken age, average daily body weight, stocking density, and environmental conditions (temperature, humidity, wind speed, and NH_3 concentration). By entering the latest farm data, users can instantly obtain mortality estimates for horizons H+1 to H+7.

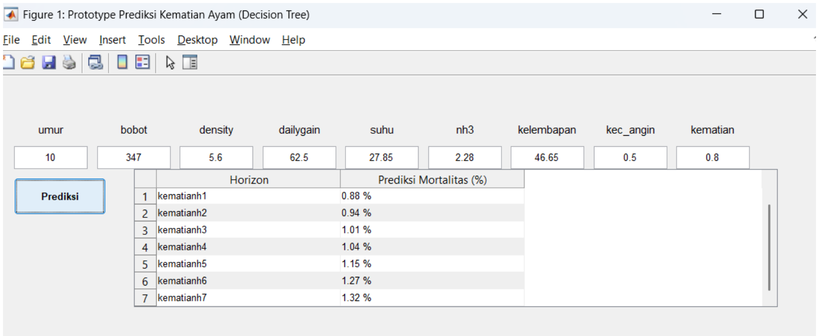

Figure 12. Display of the Prediction Prototype

The interface presents the outputs in both tabular and graphical formats, making the results more accessible to non-technical users such as farmers. An example visualization is shown in Figure 12, illustrating the initial design of the mortality prediction prototype. This visualization helps farmers understand potential mortality trends while also identifying dominant factors influencing the predictions.

Although still in the form of a research prototype, this deployment serves as an initial step toward the practical use of prediction models in broiler farm management.

Practical recommendations include:

- Farmers can utilize the system as an early warning tool to adjust ventilation, improve feed quality, and strengthen biosecurity when mortality risk is predicted to increase.
- For future development, the MATLAB GUI prototype can be expanded into a web-based or mobile application connected to IoT sensors inside poultry houses. This would allow predictions to be generated in real time and adaptively.

Thus, deployment not only serves as a technical demonstration but also provides a roadmap toward more practical, scalable, and farmer-friendly applications of broiler mortality prediction systems.

3.5 Discussion

The results in Table 6 show that both algorithms have different performance characteristics according to the prediction horizon. In the short-term horizon (H+1 and H+2), SVR produced the best performance with R^2 values of 0.84271 and 0.75982, slightly higher than Decision Tree Regression. In addition, the RMSE and CV-RMSE values of SVR were also lower, indicating that SVR predictions were more precise in the short term.

However, as the prediction horizon increased, the performance of SVR tended to decline. For example, at H+5 to H+7, the R^2 values of SVR ranged only between 0.684 and 0.684–0.706, whereas decision tree regression was relatively more stable at around 0.683–0.696. Although the R^2 values of the decision tree were not very high, their stability shows that this algorithm is more robust to long-term error accumulation.

The CV-RMSE values support these findings. At the initial horizon (H+1), the CV-RMSE of SVR was 0.43074, better than Decision Tree (0.74662). However, at the final horizon (H+7), the difference between SVR (0.59194) and Decision Tree (1.08318) became larger, indicating that Decision Tree was more affected by data variability when predictions were extended.

These findings are consistent with study (Ahmad, 2009), which reported that kernel-based methods such as SVR tend to excel in short-term predictions due to their ability to capture nonlinear patterns, but their accuracy decreases in the long term because of sensitivity to noise. Conversely, decision tree regression is easier to interpret and more stable for long-term horizons, although its initial accuracy is lower.

Overall, it can be concluded that SVR is more suitable for short-term daily mortality prediction (H+1–H+2), while decision tree regression can be used as a more stable alternative for medium- to long-term horizons. This supports the use of hybrid models or ensemble learning approaches as future research to combine the advantages of both algorithms.

Nevertheless, this study has several limitations. First, the study used a limited number of variables; for instance, not all humidity and NH_3 data were consistently available across all periods. Second, the scope of the study was limited to several poultry houses in one region, so generalizing the results to other farm contexts requires further testing. Third, the models were tested only on historical data without real-time implementation in the field.

Although SVR showed slightly better performance in the short-term horizon, deployment in this study was chosen using decision tree regression. The main consideration was the ease of interpretation and visualization of the decision tree, as well as the simplicity of integration into the MATLAB GUI prototype. With this approach, the prediction results not only provide estimated mortality values but also information on the dominant variables influencing chicken mortality, making them more useful for non-technical users such as farmers. SVR was still reported at the evaluation stage as a benchmark and may be considered for development in future web- or mobile-based systems.

4. CONCLUSION

This study successfully implemented and compared machine learning models, specifically Decision Tree Regression (DTR) and Support Vector Regression (SVR), for multi-horizon forecasting (H+1 up to H+7) of broiler mortality using real-world data from Small-Medium Farms in Sukabumi. The comparative analysis demonstrated that SVR provided superior performance for short-term prediction (H+1, with an R^2 of 0.842). In contrast, DTR exhibited greater predictive stability across the medium-to-long prediction horizons (H+5 to H+7, with R^2 values ranging from 0.683 to 0.696). Feature selection confirmed that **Age, Average Weight, Temperature, and Ammonia (NH_3)** are the dominant factors influencing mortality in this context. This research contributes a practical proof-of-concept through a MATLAB-based Graphical User Interface (GUI) prototype, which can serve as a functional early warning system to assist farmers in taking proactive measures. Future research should focus on integrating this model with real-time Internet of Things (IoT) sensor data and conducting further validation across diverse farm scales and geographical locations to enhance generalizability.

REFERENCES

Adelia, D., & others. (2024). Strategi Pengembangan Usaha Ayam Ras Petelur Di Kecamatan Dua Pitue Kabupaten Sidenreng Rappang. Universitas Hasanuddin.

Ahmad, H. A. (2009). Poultry growth modeling using neural networks and simulated data. *Journal of Applied Poultry Research*, *18*(3), 440–446. https://doi.org/10.3382/japr.2008-00064

Alves, A. A. C., Fernandes, A. F. A., Breen, V., Hawken, R., & Rosa, G. J. M. (2024). Monitoring mortality events in floor-raised broilers using machine learning algorithms trained with feeding behavior time-series data. *Computers and Electronics in Agriculture*, 224. https://doi.org/10.1016/j.compag.2024.109124

Cruz, E., Hidalgo-Rodriguez, M., Acosta-Reyes, A. M., Rangel, J. C., & Boniche, K. (2024). AI-Based Monitoring for Enhanced Poultry Flock Management. *Agriculture (Switzerland)*, 14(12). https://doi.org/10.3390/agriculture14122187

Darmawi, D. Y., Nurcahyo, G. W., & Sumijan, S. (2020). Fuzzy Sistem Fuzzy Menggunakan Metode Sugeno Dalam Akurasi Penentuan Suhu Kandang Ayam Pedaging. *Jurnal Informasi Dan Teknologi*. https://doi.org/10.37034/jidt.v3i2.95

Elghardouf, N., Lahlouh, I., Elakkary, A., & Sefiani, N. (2023). Towards modelling, and analysis of differential pressure and air velocity in a mechanical ventilation poultry house: Application for hot climates. *Heliyon*, 9(1). https://doi.org/10.1016/j.heliyon.2023.e12936

Harrison, D. L., Stanford, P. E., Sorenson, C. I., Abplanalp, H., MILLIGAN, J. L., & WINN, P. N. (1958). TURKEY COOKERY 817 The Influence of Temperature and Humidity on Broiler Performance in Environmental Chambers 12*. In Food Research* (Vol. 23).

Imam Baihaqi, D., Nur Handayani Jurusan Elektro, A., & Pujianto Jurusan Elektro, U. (2019). Perbandingan Metode Naïve Bayes Dan C4.5 Untuk Memprediksi Mortalitas Pada Peternakan Ayam Broiler. *Jurnal SIMETRIS*, *10*(1).

Jainonthee, C., Sanwisate, P., Sivapirunthep, P., Chaosap, C., Mektrirat, R., Chadsuthi, S., & Punyapornwithaya, V. (2025). Data-driven insights into pre-slaughter mortality: Machine learning for predicting high dead on arrival in meat-type ducks. *Poultry Science*, *104*(1). https://doi.org/10.1016/j.psj.2024.104648

Kader, M. S., Ahmed, F., & Akter, J. (2021). Machine learning techniques to precaution of emerging disease in the poultry industry. *2021 24th International Conference on Computer and Information Technology* (ICCIT), (pp. 1–6).

Kholil, M., Priya Waspada, H., Akhsani, R., & Komunitas Negeri Putra Sang Fajar Blitar, A. (2022). Klasifikasi Penyakit Infeksi Pada Ayam Berdasarkan Gambar Feses Menggunakan Convolutional Neural Network. *SINTECH JOURNAL*, *5*(2). https://doi.org/10.31598

Kim, D. H., Lee, Y. K., Kim, S. H., & Lee, K. W. (2021). The impact of temperature and humidity on the performance and physiology of laying hens. *Animals*, *11*(1), 1–12. https://doi.org/10.3390/ani11010056

Liu, L., Ren, M., Ren, K., Jin, Y., & Yan, M. (2020). Heat stress impacts on broiler performance: a systematic review and meta-analysis. *Poultry Science*, *99*(11), 6205–6211. https://doi.org/10.1016/j.psj.2020.08.019

Neethirajan, S. (2020). The role of sensors, big data and machine learning in modern animal farming. In *Sensing and Bio-Sensing Research* (Vol. 29). Elsevier B.V. https://doi.org/10.1016/j.sbsr.2020.100367

Pirompud, P., Sivapirunthep, P., Punyapornwithaya, V., & Chaosap, C. (2024). Application of machine learning algorithms to predict dead on arrival of broiler chickens raised without antibiotic program. *Poultry Science*, *103*(4). https://doi.org/10.1016/j.psj.2024.103504

Riber, A. B., & Wurtz, K. E. (2024). Impact of Growth Rate on the Welfare of Broilers. *In Animals* *14*(22),. Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/ani14223330

Shwetha, V., Maddodi, B. S., Laxmi, V., Kumar, A., & Shrivastava, S. (2024). Latest Trend and Challenges in Machine Learning– and Deep Learning–Based Computational Techniques in Poultry Health and Disease Management: A Review. *In Journal of Computer Networks and Communications,* (Vol. 2024). John Wiley and Sons Ltd. https://doi.org/10.1155/2024/8674250

Wei, F. X., Hu, X. F., Xu, B., Zhang, M. H., Li, S. Y., Sun, Q. Y., & Lin, P. (2015). Ammonia concentration and relative humidity in poultry houses affect the immune response of broilers. *Genetics and Molecular Research*, *14*(2), 3160–3169. https://doi.org/10.4238/2015.April.10.27

Yerpes, M., Llonch, P., & Manteca, X. (2020). Factors associated with cumulative first-week mortality in broiler chicks. *Animals*, *10*(2). https://doi.org/10.3390/ani10020310

Zabir, M., Miah, M. A., Alam, M., Bhuiyan, M. E. J., Haque, M. I., Sujan, K. M., & Mustari, A. (2021). Impacts of stocking density rates on welfare, growth, and hemato-biochemical profile in broiler chickens. *Journal of Advanced Veterinary and Animal Research*, *8*(4), 642–649. https://doi.org/10.5455/javar.2021.h556

Ziebe, S. D., Vougat Ngom, R., Akoussa, A. M. M., Bogning, H. P., & Zangue, H. A. (2025). Impact of Biosecurity on Production Performance and Antimicrobial Usage in Broiler Farms in Cameroon. *Animals*, *15*(12). https://doi.org/10.3390/ani15121771