ICIC Express Letters
Volume **19**, Number **12**, December **2025**
 ISSN 1881-803X
pp. 1313–1322

FAT FILLED POWDER QUALITY EVALUATION IN IMBALANCED DATA CONDITION FOR DAIRY PRODUCTS BASED INDUSTRY

Oki Pratama and Muhammad Asrol*

Industrial Engineering Department, BINUS Graduate Program – Master of Industrial Engineering
Bina Nusantara University
Jl. K. H. Syahdan No. 9, Kemanggisan, Palmerah, Jakarta 11480, Indonesia
oki.pratama@binus.ac.id; *Corresponding author: muhammad.asrol@binus.edu

Received December 2024; accepted March 2025

Abstract. *The dairy processing industry has a strategic role in fulfilling people's nutrition. However, traditional methods of determining product quality take a long time, which can lead to product damage and financial losses. Therefore, an intelligent quality control system is required to improve efficiency and accuracy in quality determination. This research aims to analyze quality criteria and design a fat filled powder quality assessment model using a machine learning approach. Machine learning approach with Decision Tree, AdaBoost, Gradient Boosting and Random Forest are proposed to model the quality assessment. To validate model performance, a cross validation was proposed. The results showed that all proposed model has good performance and is validated using cross-validation. The Decision Tree model showed the best results compared to Gradient Boosting, Random Forest, and AdaBoost, with an area under the curve (AUC) of 0.989, and accuracy, F1 score, recall, and precision of 0.994, and matthews correlation coefficient (MCC) of 0.979. Machine learning-based approaches, particularly Decision Tree models with cross validation, were shown to improve the effectiveness of fat filled powder quality control in the dairy processing industry. For further industrial applications, a system installation for model implementation is also provided.*
Keywords: Fat filled powder, Quality control, Machine learning, Cross validation, Decision Tree

1. **Introduction.** Competition in the industrial world is becoming increasingly intense, requiring businesses to enhance product quality and meet consumer expectations. As a result, production processes and quality control play a crucial role in ensuring products meet standards and consumer demands [1,2]. The dairy processing industry has an important and strategic role in efforts to provide and fulfill community nutrition. According to the Indonesian Central Bureau of Statistics, the average milk consumption in Indonesia is only 16.27 kg/capita/year, while only about 22% of milk raw materials are supplied domestically and the other 78% are imported [3]. Fat filled powders are an alternative raw material used for the dairy industry [4]. Fat filled powders are made from skim milk with added vegetable fat as a milk fat substitute, requiring homogenization before spray drying for effective emulsification [5]. In addition, the protein and fat content of milk affects the quality of dairy products [6].

These ingredients may degrade product quality, failing to meet company and national standards, which reduces sales and consumer interest. This increases industry competition to produce high-quality dairy products amid increasingly complex nutritional requirements [7]. The dairy industry faces quality uncertainty in continuous production, requiring effective control. Manual monitoring of fat filled powder quality is error-prone and inefficient, relying on a single, unreliable indicator. Multi-feature analysis is needed to enhance accuracy, efficiency, and quality evaluation using advanced approaches. Along

DOI: 10.24507/icicel.19.12.1313

with the development of science and technology, it is possible to apply smart quality evaluation models. The smart quality evaluation model can integrate machine learning and quality assurance knowledge in improving product quality. Some previous studies have used machine learning approaches in smart quality evaluation. For example, [8] used random forest, K-nearest neighbors (KNN), and neural network (NN) for milk quality classification. [9] used partial least squares (PLS) and artificial neural network (ANN) to predict Lactoferrin content. Furthermore, [10] utilized Decision Tree, artificial neural network and AdaBoost, while [11] applied artificial neural network, Decision Tree, Gradient Boosting, and vector space model (VSM) to analyzing fat filled powder quality in synthetic milk within a continuous production scheme. It introduces a smart monitoring system for real-time assessment and prioritizes machine learning and artificial intelligence over deep learning, focusing on bagging and boosting for classification.

A challenge to this research is the occurrence of imbalanced data, a condition where the distribution of classes in a data set is uneven, with one or more classes having a much smaller number of samples than other classes [12]. This condition often occurs in binary and multi-class classification problems [13] and it may affect the machine learning algorithms to be biased towards the majority class (the class with the largest number of samples). This condition led the accuracy degradation for minority class [14]. Overcoming imbalanced data in machine learning can be done through several approaches, such as data resampling (oversampling with Synthetic Minority Over Sampling Technique (SMOTE) or under sampling to balance the class distribution), weight adjustment to give more attention to minority classes, and selection of imbalance-resistant algorithms such as Random Forest or Gradient Boosting. In addition, the use of appropriate evaluation metrics such as F1 score, area under the curve (AUC), and matthews correlation coefficient (MCC) can provide a more accurate picture of the model's performance in handling imbalanced data [15,16].

This research seeks to contribute to enhancing fat filled powder quality monitoring with machine learning model development. This research also contributes to developing an accurate model of quality prediction with challenges in data imbalanced condition. Decision Tree-based model and its optimized model are also evaluated to test the model accuracy in monitoring fat filled powder quality in imbalanced data history condition. By applying machine learning considering imbalanced data condition for quality evaluation, it becomes faster, automated, and consistent, thus reducing laboratory waiting time and improving production efficiency.

This research aims to develop a machine learning model to evaluate fat filled powder quality in imbalanced dataset conditions. A Decision Tree-based model and its ensemble model are provided to develop the model to evaluate the model accuracy in imbalanced data histories condition.

This paper is organized as follows. In Section 2, related work is provided to elaborate previous research related to the proposed model and define the research gap and contribution. In Section 3, method is provided to explain the step-by-step research stage in model development and evaluation. In Section 4, results and discussion are provided to explain research result and the accurate model for fat filled powder quality evaluation. Finally, in Section 5, the conclusion and recommendations are elaborated.

2. **Related Work.** Milk is a perishable product affected by uncertain substances, including fat filled powder. To the best of the authors' knowledge, limited studies exist on evaluating synthetic milk quality using a machine learning-based smart evaluation system and quality control under continuous production, especially in Indonesia. Automation helps the food sector effectively manage product quality for human consumption, enabling food fraud detection systems for comprehensive integrity and content analysis [17]. These employ deep learning, machine learning, or artificial intelligence algorithm learning [18].

For example, Dabija et al. [19] developed a milk quality monitoring system using analytical methods. Neto et al. [20] utilized deep and ensemble learning to observe the quality of milk additives. Soyeurt et al. [9] used infrared spectra in predicting milk quality and combined with machine learning. Machine learning was also utilized by León and Ossa [8] for dairy product classification. If you pay attention, the study of milk quality in previous research is also diverse, namely the use of analytical methods, machine learning and infrared spectra. According to Jiménez-Carvelo et al. [21], product quality would be best monitored using a machine learning approach. This opens up wider research opportunities on milk quality at the micro level of quality.

Although various previous studies have explored milk quality monitoring using analytical methods, deep learning, ensemble learning, infrared spectra, and machine learning, studies related to Fat Filled Powder (FFP) quality evaluation in continuous milk production schemes are still limited, especially in Indonesia. However, no research has specifically addressed the evaluation of FFP quality in sustainable dairy production with a multi-metric-based machine learning approach. Therefore, this study proposes an FFP quality evaluation model using Decision Tree, Gradient Boosting, Random Forest, as well as methods without boosting and ensemble, with the selection of the best model based on accuracy, AUC, precision, F1 score, recall, and MCC. This opens up further research opportunities in the development of a more precise and efficient intelligent monitoring system for the dairy industry.

3. **Method.** The stages and flowchart as well as the success indicators of the research can be seen in Figure 1. This research will be completed through six main stages, where the initial three stages have been completed and are being worked on. Thus, at this stage only the next three stages will be completed. In the first stage, the urgency and motivation for the research were finalized. Indicators and outputs obtained are the knowledge of the gap

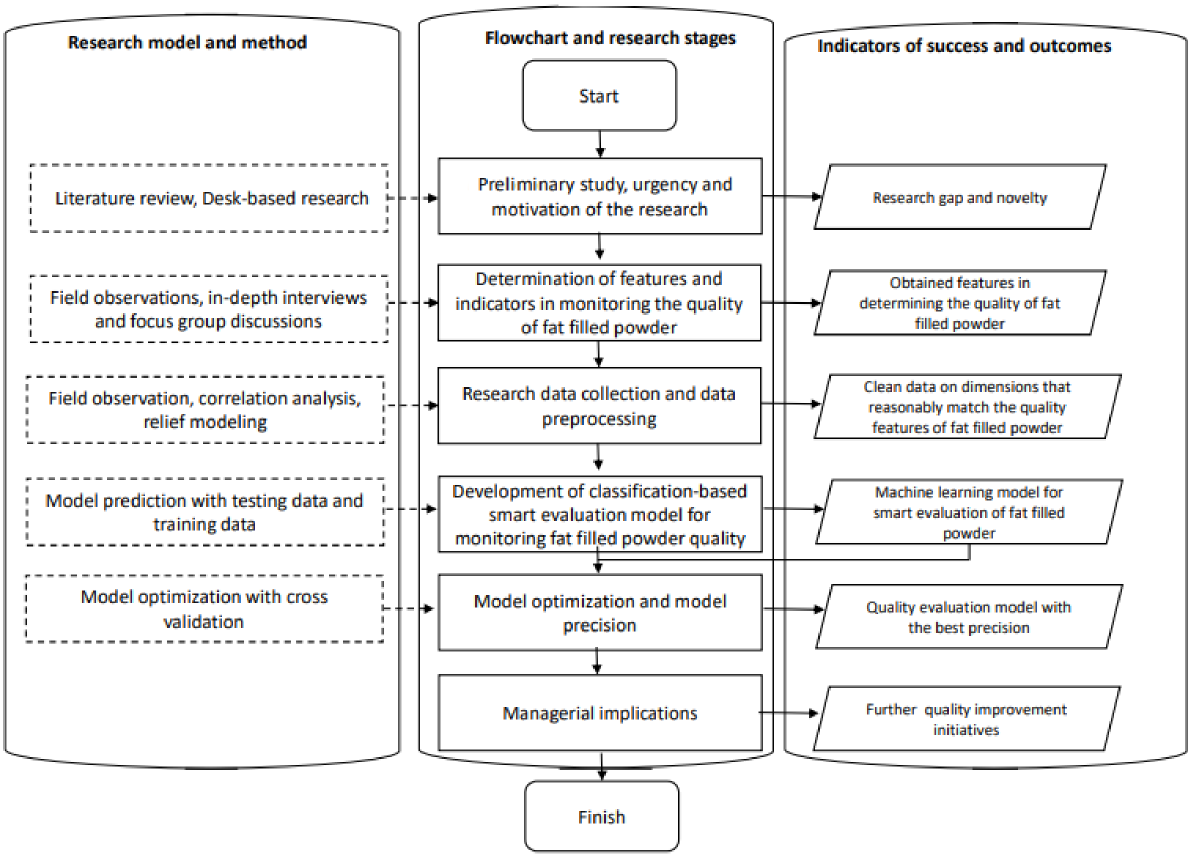

Figure 1. Research stages

and novelty of the research in terms of monitoring fat filled powder quality in a continuous production system through machine learning. In the second and third stages, the features and main indicators in fat filled quality monitoring were determined. This dataset consists of seven features and one target variable, the quality of fat filled powder. The data has dimensions of 1525×7 and is in a clean state. The data has gone through preprocessing and data normalization, transforming the raw data obtained from field observations. This stage has obtained the main features in system development which can be seen in Table 3.

In the fourth stage, we will develop a smart evaluation model using Decision Tree, Gradient Boosting, Random Forest and AdaBoost models. Firstly, we prepare the testing and training data for model development, as also provided in [22]. While in the fifth stage, model optimization is carried out with cross validation. In order to develop the model, this research employs the 10-fold cross-validation technique, similar to a study by [7]. This research uses ensemble learning with bagging (Decision Tree, Random Forest) and boosting (Gradient Boosting, AdaBoost) models according to the algorithms in Tables 1 and 2. The primary objective of employing bagging and boosting techniques is to mitigate the issues of overfitting and underfitting, thereby yielding outcomes with reduced bias and variance. Bagging employs a bootstrapping sampling technique, while boosting utilizes reweighting sampling based on the data distribution obtained from previous models [7].

TABLE 1. Bagging algorithm [7]

Bagging Algorithm
Input: Dataset $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$;
Base learners T
Base learning algorithm $\pounds$
For $t = 1, \ldots, T$:
$h_t = \pounds(D, D_{bs})$; D_{bs} as boostrap sampling distribution
End
Output: $H(x) = \text{argmax} \sum_{t=1}^{T} \Pi(h_t(\boldsymbol{x}) = y)$

TABLE 2. Boosting algorithm [7]

Boosting Algorithm
Input: sample distribution D;
Base learners T;
Base learning algorithm $\pounds$;
$D_1 = D$; initialize distribution
For $t = 1, \ldots, T$:
$h_t = \pounds(D_t)$; Train a weak learner from distribution D_t
$\epsilon_t = PP_{\boldsymbol{x} \sim D_t}(h_t(\boldsymbol{x}) \neq f(\boldsymbol{x}))$; Evaluate the error of h_t
$D_{t+1} = \textit{Adjust\_Distribution}\ (D_t, \epsilon_t)$
End
Output: $H(x) = \text{Combine\_Outputs}(\{h_1(\boldsymbol{x}), \ldots, h_t(\boldsymbol{x})\})$

Confusion matrix is proposed to evaluate the model performance. This involves comparing the predicted outcomes of the model with the actual data to assess its performance. In Table 4, the correspondence between the predicted class (i) and the actual class (j) is is analyzed to determine the agreement between the model's predictions and the actual class labels. Suppose that x_{11}, x_{21}, x_{12}, and x_{22} represent number of true positives, false positives, false negatives, and true negatives, respectively. Based on these values, the model's performance can be calculated. After obtaining the values from the confusion

Table 3. Key features in determining the quality of fat filled powder

No	Attributes	Data type	Role	Unit	Max	Min	Average	St. Dev.
1	Fat	Numeric	Feature	%	34.6	27.0	32.70	0.58
2	Protein	Numeric	Feature	%	20.7	18.5	19.20	0.29
3	Lactose	Numeric	Feature	%	42.7	37.7	39.10	0.34
4	Ash	Numeric	Feature	%	6.3	4.1	5.20	0.30
5	Moisture	Numeric	Feature	%	2.7	1.4	2.30	0.14
6	Bulk density	Numeric	Feature	g/ml	0.60	0.50	0.50	0.01
7	Acidity	Numeric	Feature	%	12.1	8.8	10.50	0.47
8	FFP quality	Categorical	Target	(Good/Bad)				

Table 4. Confusion matrix

Confusion matrix		Predicted class (i)	
		Bad quality	Good quality
Actual class (j)	Bad quality	x_{11}	x_{21}
	Good quality	x_{12}	x_{22}

matrix, the model's performance can be evaluated using accuracy, precision, F1 score, recall, and AUC, as shown in Equations (1)-(4).

$$Accuracy = \frac{x_{11} + x_{22}}{\sum_{i=1}^{2} \sum_{j=1}^{2} x_{i,j}} \tag{1}$$

$$Precision = \frac{x_{11}}{x_{11} + x_{22}} \tag{2}$$

$$Recall = \frac{x_{11}}{x_{11} + x_{21}} \tag{3}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{4}$$

4. **Main Results.** The ensemble method combines multiple machine learning classifiers to enhance performance beyond a single model. This study employs two ensemble models, bagging and boosting, which have been widely validated for classification effectiveness. The parameters used in the ensemble method model are detailed in Table 5.

Table 5. Model parameters

No	Ensemble method	Model	Parameter	Value
1	Bagging	Decision Tree	Min. Number of instances in leaves	4
			Do not split subsets smaller than	10
			Limit the maximal tree dept to	100
2	Boosting	Gradient Boosting	Number of trees	100
			Learning rate	0.100
			Limit the depth of individual tree	3
			Fraction of training	1.00
3	Bagging	Random Forest	Number of trees	14
			Number of attributes each split	5
			Allowable limit depth trees	No
4	Boosting	AdaBoost	Number of estimators	50
			Learning rate	1.00

4.1. **Designing a quality evaluation model with Decision Tree.** The performance of the FFP quality evaluation model design of the Decision Tree model is further illustrated through Table 6. The model was assessed using a confusion matrix with six metrics: AUC, accuracy, F1 score, precision, recall, and MCC. The development of this model was done with testing and training data and then with cross validation using 10 times cross validation.

TABLE 6. Model evaluation of Decision Tree

Metode	Model	AUC	CA	F1	Prec	Rec	MCC
Data testing & training	Tree	0.990	0.990	0.990	0.990	0.990	0.966
Cross validation	Tree	0.989	0.994	0.994	0.994	0.994	0.979

In accordance with Table 6, it is concluded that the performance of the Decision tree model with cross validation is superior compared to using training data and testing data. Confusion matrix of the Decision Tree model can be seen in Table 7 and Table 8.

TABLE 7. Confusion matrix of Decision Tree data testing and training

Confusion matrix		Predicted class (i)		Total
		Bad quality	Good quality	
Actual class (j)	Bad quality	557	13	570
	Good quality	19	2621	2640
Total		576	2634	3210

TABLE 8. Confusion matrix of Decision Tree cross validation

Confusion matrix		Predicted class (i)		Total
		Bad quality	Good quality	
Actual class (j)	Bad quality	231	5	236
	Good quality	3	1133	1136
Total		234	1138	1372

Meanwhile, the ROC curve of the Decision Tree on model prediction using testing data and training data shows that the AUC value = 0.500 indicates that the model has good performance, while the ROC curve on the model with cross validation shows that the AUC value = 0.600 indicates that the model has good performance. Model results with testing and training data show that the Matthew's correlation coefficient (MCC) accuracy value is superior with a value of 0.966, compared to cross validation of 0.979. The conclusion on both methods in determining the performance of the model with this Decision Tree is that the method with cross validation is better and outperforms the existing matrix value compared to the training and testing data.

4.2. **Quality evaluation model design with Gradient Boosting.** The performance of the FFP quality evaluation model design of the Gradient Boosting model is further illustrated through Table 9. The model was assessed using a confusion matrix with six metrics: AUC, accuracy, F1 score, precision, recall, and MCC. The development of this model was done with testing and training data and then with cross validation using 10 times cross validation.

In accordance with Table 9, it can be concluded that the performance of the Gradient Boosting model with the cross validation method shows superior performance compared to using testing data and training data. The Gradient Boosting model confusion matrix can be seen in Table 10 and Table 11.

Table 9. Model evaluation of Gradient Boosting

Metode	Model	AUC	CA	F1	Prec	Rec	MCC
Data testing & training	Gradient Boosting	0.987	0.993	0.993	0.993	0.993	0.975
Cross validation	Gradient Boosting	0.995	0.993	0.993	0.993	0.993	0.974

Table 10. Confusion matrix of Gradient Boosting data testing and training

Confusion matrix		Predicted class (i)		Total
		Bad quality	Good quality	
Actual class (j)	Bad quality	556	14	570
	Good quality	9	2631	2640
Total		565	2645	3210

Table 11. Confusion matrix of Gradient Boosting cross validation

Confusion matrix		Predicted class (i)		Total
		Bad quality	Good quality	
Actual class (j)	Bad quality	229	7	236
	Good quality	3	1133	1136
Total		232	1140	1372

The ROC curve on model prediction using testing data and training data shows AUC = 0.637 indicating that the model has good performance, while the ROC curve on cross validation shows that the gradient boosting model AUC = 0.378 indicates the model has poor performance. The model results show that the MCC or Matthew's correlation coefficient (MCC) accuracy value shows the testing data is superior with a value of 0.975, while with cross validation it is 0.974. The conclusion on both methods in determining the performance of the Gradient Boosting model is that the cross validation method is better and outperforms the existing matrix value compared to the training and testing data.

4.3. **Designing quality evaluation model with Random Forest.** The performance of the FFP quality evaluation model design of the Random Forest model is further illustrated through Table 12. The model was assessed using a confusion matrix with six metrics: AUC, accuracy, F1 score, precision, recall, and MCC. The development of this model was done with testing and training data and then with cross validation using 10 times cross validation.

Table 12. Model evaluation of Random Forest

Metode	Model	AUC	CA	F1	Prec	Rec	MCC
Data testing & training	Random Forest	0.995	0.992	0.992	0.992	0.992	0.971
Cross validation	Random Forest	0.992	0.992	0.992	0.992	0.992	0.972

In accordance with Table 12, it can be concluded that the performance of the Random Forest model with the cross validation method shows superior performance compared to determining training data and testing data. The confusion matrix of the Random Forest model can be seen in Table 13 and Table 14.

The ROC curve on the model using testing data and training data shows that the random forest model has AUC = 0.500 which indicates that the model is not better than random guesses, while the ROC curve on the model using cross validation shows that the Random Forest model has AUC = 0.429 which indicates that the model is not better

TABLE 13. Confusion matrix of Random Forest data training and testing

Confusion matrix		Predicted class (i)		Total
		Bad quality	Good quality	
Actual class (j)	Bad quality	549	21	570
	Good quality	6	2643	2649
Total		555	2664	3219

TABLE 14. Confusion matrix of Random Forest cross validation

Confusion matrix		Predicted class (i)		Total
		Bad quality	Good quality	
Actual class (j)	Bad quality	226	10	236
	Good quality	1	1135	1136
Total		227	1145	1372

than random guesses. The results of the Random Forest model show that the accuracy value of Matthew's correlation coefficient (MCC) shows that cross validation is superior with a value of 0.972 compared to the value of testing data and training data of 0.971. The conclusion on these two methods in determining the performance of the model is that the method with cross validation is better and outperforms the existing matrix value compared to the training and testing data.

4.4. **Quality evaluation model design with AdaBoost.** The FFP quality evaluation AdaBoost model design performance of the AdaBoost model is further illustrated through Table 15. The model was assessed using a confusion matrix with six metrics: AUC, accuracy, F1 score, precision, recall, and MCC. The development of this model was done with testing and training data and then with cross validation using 10 times cross validation.

TABLE 15. Model evaluation of AdaBoost

Metode	Model	AUC	CA	F1	Prec	Rec	MCC
Data testing & training	AdaBoost	0.986	0.992	0.992	0.992	0.992	0.971
Cross validation	AdaBoost	0.986	0.993	0.993	0.993	0.993	0.974

In accordance with Table 15, it is concluded that the performance of the AdaBoost model with the cross-validation method shows superior performance compared to using training data and testing data. The confusion matrix of the AdaBoost model can be seen in Table 16 and Table 17.

TABLE 16. Confusion matrix of AdaBoost data testing & training

Confusion matrix		Predicted class (i)		Total
		Bad quality	Good quality	
Actual class (j)	Bad quality	556	14	570
	Good quality	11	2629	2640
Total		567	2643	3210

The ROC curve on the model using testing data and training data as well as cross validation shows that the AdaBoost model has perfect performance because it has AUC = 1. The results of the AdaBoost model show that the accuracy value of Matthew's correlation coefficient (MCC) cross validation is superior with a value of 0.974 compared to testing and training data which is 0.971. The conclusion on these two methods is that the method with cross validation is better and outperforms the existing matrix value compared to the training and testing data.

TABLE 17. Confusion matrix of AdaBoost cross validation

Confusion matrix		Predicted class (i)		Total
		Bad quality	Good quality	
Actual class (j)	Bad quality	230	6	236
	Good quality	4	1132	1136
Total		234	1138	1372

4.5. **Managerial implications and research limitations.** The managerial implication of the quality evaluation model is to eliminate the conventional way of inputting data in excel in determining the quality of fat filled powder. With the implementation of the Decision Tree model, product quality evaluation becomes faster, automatic, and consistent. This can reduce waiting time in the laboratory or manual checking, allowing the company to increase production efficiency. Cost savings can also occur from reducing waste of products that do not meet standards. Figure 2 shows an illustration of the fat filled powder (FFP) quality evaluation process using machine learning, where data is analyzed, models are trained, predictions are made, and the results are evaluated to determine product quality (good/bad).

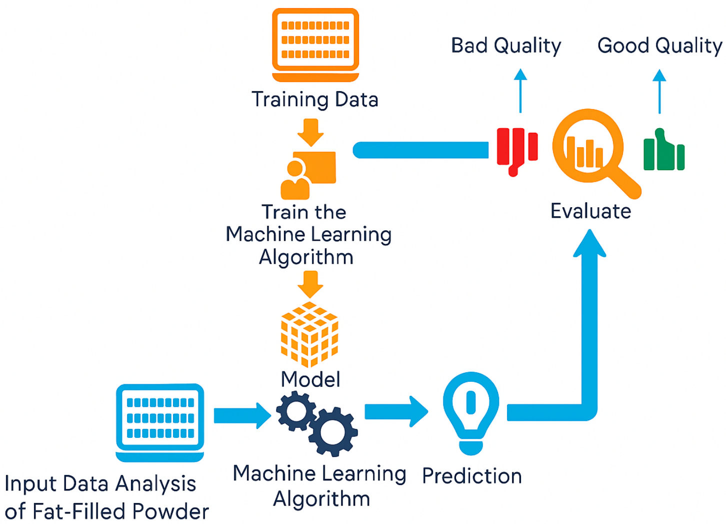

FIGURE 2. Illustration of the implications of the quality evaluation model

The main limitation of this study is the potential changes in food industry regulations and quality standards. As these standards evolve over time, the model may require adjustments to align with updated quality parameters.

5. **Conclusions.** Fat filled powder testing criteria consider food quality safety standards, these criteria are fat, protein, lactose, ash, moisture, bulk density, and acidity. Machine learning is used in designing models in determining quality, namely by comparing the performance of determining model performance using testing and training data and by cross validation. In this study, the results show that the cross-validation method has superior results in terms of matrix values AUC, F1 score, accuracy, recall, precision and MCC. By using the cross-validation method, it is found that the Decision Tree model has superior performance compared to the AdaBoost, Gradient Boosting, and Random Forest models. For further research potential, it can explore the use of deep learning to improve the accuracy of fat filled powder quality evaluation and develop the implementation of IoT-based monitoring system for real-time quality monitoring.

REFERENCES

[1] B. V. R. Guzmán and O. F. C. Domínguez, Benefits on productivity indicators after quality management implementation: Evidence in the dairy industry, *Int. J. Product. Perform. Manag.*, vol.72, no.8, pp.2349-2375, 2023.

[2] S. O. Akhmetova and M. S. Suleimenova, Quality management system for improvement of quality and efficiency of food production: Case of dairy products enterprise, *Entrep. Sustain. Issues*, vol.6, no.1, pp.289-310, 2018.

[3] Center for Agricultural Data and Information Systems, *Livestock Commodity Outlook: Milk*, Jakarta, 2022.

[4] M. H. Patil et al., Energy usage in the manufacture of dairy powders: Advances in conventional processing and disruptive technologies, *Drying Technology*, vol.39, no.11, pp.1595-1613, 2021.

[5] E. W. Finnegan, M. S. Mahomud, E. G. Murphy and J. A. O'Mahony, The influence of pre-heat treatment of skim milk on key quality attributes of fat filled milk powder made therefrom, *Int. J. Dairy Technol.*, vol.74, no.2, pp.404-413, 2021.

[6] T. I. Purwantiningsih, M. A. B. Bria and K. W. Kia, Levels protein and fat of yoghurt made of different types and number of cultures, *J. Trop. Anim. Sci. Technol.*, vol.4, no.1, pp.66-73, 2022.

[7] M. Asrol, S. Wahyudi, Suharjito, C. Harito, D. N. Utama and M. Syafrudin, Improving supplier evaluation model using ensemble method-machine learning for food industry, *Procedia Comput. Sci.*, vol.227, pp.307-315, https://doi.org/10.1016/j.procs.2023.10.529, 2023.

[8] M. León and D. Ossa, Machine learning applied to milk sample classification, *Proc. of the 1st Australian International Conference on Industrial Engineering and Operations Management*, pp.2390-2398, 2022.

[9] H. Soyeurt, C. Grelet, S. McParland, M. Calmels, M. Coffey, A. Tedde et al., A comparison of 4 different machine learning algorithms to predict lactoferrin content in bovine milk from mid-infrared spectra, *J. Dairy Sci.*, vol.103, no.12, pp.11585-11596, 2020.

[10] A. Ç. Elik, Using machine learning algorithms to detect milk quality, *Eurasian J. Food Sci. Technol.*, vol.6, pp.76-87, 2022.

[11] S. Wahyudi and M. Asrol, Designing a supplier evaluation model in the cheese industry using hybrid method, *Acad. J. Manuf. Eng.*, vol.20, no.2, pp.27-35, 2022.

[12] H. Wang et al., Prediction of flood risk levels of urban flooded points though using machine learning with unbalanced data, *Journal of Hydrology*, vol.630, 130742, 2024.

[13] M. Bacevicius and A. Paulauskaite-Taraseviciene, Machine learning algorithms for raw and unbalanced intrusion detection data in a multi-class classification problem, *Applied Sciences*, vol.13, no.12, 7328, 2023.

[14] S. Chatterjee and Y.-C. Byun, Highly imbalanced fault classification of wind turbines using data resampling and hybrid ensemble method approach, *Engineering Applications of Artificial Intelligence*, vol.126, 107104, 2023.

[15] A. Fernández, S. García, M. Galar, R. C. Prati, B. Krawczyk and F. Herrera, *Learning from Imbalanced Data Sets*, Springer, 2018.

[16] M. Buda, A. Maki and M. A. Mazurowski, A systematic study of the class imbalance problem in convolutional neural networks, *Neural Networks*, vol.106, pp.249-259, 2018.

[17] C. Mengucci, D. Rabiti, E. Urbinati, G. Picone, R. Romano, A. Aiello, P. Ferranti and F. Capozzi, Spotting frozen curd in PDO Buffalo Mozzarella cheese through insights on its supramolecular structure acquired by 1H TD-NMR relaxation experiments, *Applied Sciences*, 2021.

[18] R. Gupta et al., Artificial intelligence to deep learning: Machine intelligence approach for drug discovery, *Molecular Diversity*, vol.25, pp.1315-1360, 2021.

[19] A. I. Dabija, S. Mironeasa, M. Oroian and I. Sion, Study concerning milk quality – Raw material for dairy industry, *ETP Int. J. Food Eng.*, pp.14-21, 2018.

[20] H. A. Neto, W. L. F. Tavares, D. C. S. Z. Ribeiro, R. C. O. Alves, L. M. Fonseca and S. V. A. Campos, On the utilization of deep and ensemble learning to detect milk adulteration, *BioData Min.*, vol.12, no.1, pp.1-13, 2019.

[21] A. M. Jiménez-Carvelo, A. González-Casado, M. G. Bagur-González and L. Cuadros-Rodríguez, Alternative data mining/machine learning methods for the analytical evaluation of food quality and authenticity – A review, *Food Res. Int.*, vol.122, pp.25-39, 2019.

[22] M. K. Uçar et al., The effect of training and testing process on machine learning in biomedical datasets, *Mathematical Problems in Engineering*, 2836236, 2020.