G-Tech : Jurnal Teknologi Terapan
Vol. 10, No. 3, July 2026, p. 967-973
E-ISSN: 2623-064X | P-ISSN: 2580-8737

Silk Fabric Quality Prediction Using Random Forest and Simulated Physical Material Feature Data

Moh. Zainuddin[1]✉, Rachmat[2], Muh. Rafli R.[3]

[1] Informatics Enginering Study Program, Faculty of Technology and Design, Institut Teknologi dan Bisnis Asia Malang, Indonesia
[2] Informatics Enginering Study Program, Faculty of Enginering, Universitas Pejuang Republik Indonesia, Indonesia
[3] Information System Study Program, Faculty of Computer Science, Universitas DIPA Makassar, Indonesia

✉ **Corresponding Author** : Moh. Zainuddin (e-mail: mzein@asia.ac.id)

Article Information

Article History

Received : May 01, 2026
Revised : May 11, 2026
Published : July 02, 2026

Keywords:

Data Mining;
Feature Importance;
Machine Learning;
Physical Material Features;
Quality Control.

ABSTRACT

This study develops an interpretable Random Forest workflow to predict silk fabric quality from simulated physical material feature data. The model uses 150 controlled observations with eight predictors: fabric thickness, fabric weight, thread density, tensile strength, elasticity, luster, softness, and surface defects. The simulated values were generated with class-specific truncated normal distributions based on measurable textile properties, standard testing variables, and domain assumptions. Data were evaluated using a stratified 80:20 holdout, stratified 5-fold cross-validation, weighted accuracy-related metrics, and confusion matrix analysis. Random Forest achieved 0.87 holdout accuracy, 0.88 precision, 0.87 recall, and 0.87 F1-score. It showed competitive performance against KNN, Decision Tree, SVM, ANN, and XGBoost while providing direct feature-importance interpretation. Surface defects, tensile strength, and thread density were the most influential predictors. The main contribution is a reproducible simulation and interpretation framework for early-stage silk fabric quality prediction. Because the dataset is simulated, the results demonstrate method feasibility and require validation with real textile production or laboratory data.

INTRODUCTION

Silk fabric has high commercial value because buyers expect a smooth surface, stable structure, natural luster, adequate strength, and tactile comfort. Quality failure creates direct technical and economic loss. Fabric rolls with visible defects can be downgraded, reworked, rejected, or sold as lower-grade products. Song et al. (2021) reported that second-class textile fabrics can lose 45-65% of their price compared with first-class fabrics, while manual inspection accuracy in textile factories can reach only about 70%. Carrilho et al. (2024) also noted that human visual inspection remains vulnerable to fatigue, inconsistency, and subjective judgment. These findings show that fabric quality control needs more consistent decision-support methods.

Recent textile inspection research has focused strongly on image-based defect detection. Deep learning and computer vision methods can detect holes, stains, broken yarn, color irregularity, and other visible defects from fabric images (Li et al., 2021). More recent studies continue this direction by improving automated inspection systems for industrial environments, including acoustic-panel manufacturing and real-time vision-based defect detection (Cheung & Yang, 2024; Lin et al., 2024). These systems are useful when the objective is surface-defect localization. However, they require cameras, lighting control, image acquisition, and image-processing infrastructure. These

This is an open access article under the CC BY license

requirements can limit use in small laboratories, small enterprises, or early-stage quality-control settings.

Another research stream uses numerical material features to predict fabric properties and quality. Physical variables such as thickness, mass per unit area, thread density, tensile strength, elasticity, luster, and softness already appear in material testing and product grading. Tu et al. (2024) reviewed AI-driven prediction of fabric properties and handfeel and found that mechanical, structural, and tactile features can support fabric-property prediction. The same review also emphasized that dataset diversity and model generalization remain unresolved issues. This condition creates space for a transparent preliminary model that uses numerical physical features before real industrial data are available.

This study uses Random Forest as the main algorithm for three reasons. First, the dataset is small and tabular. Tree-based models remain strong baselines for tabular data because they can capture nonlinear relationships and feature interactions without strict distributional assumptions (Grinsztajn et al., 2022). Second, Random Forest reduces the instability of a single decision tree through bootstrap aggregation and random feature selection (Breiman, 2001). Third, the study needs both prediction and interpretation. Random Forest can rank physical variables through feature importance, which helps identify which properties deserve quality-control attention. SVM, ANN, and XGBoost can provide competitive prediction, but they require additional tuning or explainability tools to produce similarly direct variable-level interpretation.

The research gap is the limited availability of a reproducible and interpretable workflow for silk fabric quality prediction using simulated physical material feature data. Existing studies often emphasize image-based defect detection or real-material testing, while the early modeling stage receives less attention when real silk fabric datasets are not yet available. The contribution of this study is a controlled simulation protocol that connects physical testing variables, class-based assumptions, model comparison, and Random Forest feature-importance analysis. The objective is to predict silk fabric quality classes from simulated physical material features and identify the dominant variables in the simulated classification process.

RESEARCH METHODS

This study used a quantitative data-mining design with a controlled simulated dataset. The dataset contained 150 simulated silk fabric observations. Each observation had eight input features and one target label. The target label consisted of three classes: low, medium, and high quality. The class distribution was 45 low-quality samples, 60 medium-quality samples, and 45 high-quality samples. This distribution created a near-balanced multiclass problem while allowing medium-quality fabric to appear more frequently, as often occurs in production grading.

The research material was the simulated tabular dataset. No physical silk fabric specimens were tested. The selected variables represent measurable physical, mechanical, structural, visual, and tactile dimensions of fabric quality. Fabric thickness and fabric weight follow standard textile measurement concepts, including ISO 5084 for textile thickness and ISO 3801 for mass per unit area. Tensile strength and elongation also reflect established fabric testing concepts, including ASTM D5035-11(2024) for strip-test breaking force and elongation. These standards justify the relevance of the variables. They do not define the simulated class thresholds used in this study.

Table 1. Input features used for silk fabric quality prediction

No.	Feature	Unit/Scale	Description
1	Fabric thickness	Mm	Textile thickness measured under specified pressure
2	Fabric weight	g/m2	Mass per unit area
3	Thread density	thread/inch	Compactness of the fabric structure
4	Tensile strength	N	Resistance to pulling force
5	Elasticity	%	Stretch response of the fabric

No.	Feature	Unit/Scale	Description
6	Luster	1-10 scale	Visual brightness of the surface
7	Softness	1-10 scale	Tactile comfort
8	Surface defects	count	Number of visible surface defects

The simulation generated feature values using class-specific truncated normal distributions. Each feature had a mean, standard deviation, minimum value, and maximum value for low, medium, and high quality classes. Values outside the defined boundaries were truncated to avoid physically implausible data. Surface defects were rounded to integer counts. Luster and softness used a 1-10 scale. The random seed was fixed at 42 to support reproducibility.

The simulation followed explicit quality assumptions. High-quality fabric was assigned fewer surface defects, higher tensile strength, higher thread density, better luster, and better softness. Medium-quality fabric was assigned intermediate values with overlap toward both low and high classes. Low-quality fabric was assigned more defects and weaker mechanical or structural values. The ranges deliberately overlap so the classes are not perfectly separable. This design creates boundary cases that resemble adjacent grading uncertainty. However, the values remain hypothetical. They are not industrial thresholds and do not prove real-world silk fabric behavior.

The use of simulated data has direct implications for validity. Internal validity is limited to the assumptions embedded in the data-generation process. The model can only show whether the selected algorithm can recover the simulated relationship between physical features and quality labels. External validity has not been established because no real production batches, laboratory measurements, inspector labels, or factory grading standards were used. Therefore, the results must be read as a methodological demonstration, not as an industrial benchmark.

Table 2. Simulation parameters for controlled data generation

Feature	Low	Medium	High
Fabric thickness	0.18 ± 0.04 [0.1-0.3]	0.2 ± 0.04 [0.12-0.32]	0.22 ± 0.04 [0.14-0.34]
Fabric weight	54 ± 14 [28-88]	62 ± 15 [35-96]	70 ± 15 [42-110]
Thread density	76 ± 15 [55-125]	98 ± 17 [65-135]	122 ± 17 [78-150]
Tensile strength	62 ± 18 [25-115]	82 ± 20 [40-130]	102 ± 20 [55-155]
Elasticity	13 ± 5 [3-30]	15 ± 5 [5-30]	17 ± 5 [7-32]
Luster	5.2 ± 1.6 [1-8]	6.1 ± 1.5 [3-10]	7.0 ± 1.5 [4-10]
Softness	5.2 ± 1.7 [1-8]	6.1 ± 1.6 [3-10]	7.0 ± 1.5 [4-10]
Surface defects	7.5 ± 1.5 [1-14]	4.2 ± 1.5 [0-10]	1.5 ± 1.5 [0-8]

Pre-processing involved five steps. First, all feature columns were checked for numeric format consistency. Second, the dataset was checked for missing values and duplicate rows. Third, surface-defect values were converted into integer counts, while luster and softness values were restricted to the 1-10 scale. Fourth, the target label was encoded into three quality classes. Fifth, features were standardized only for scale-sensitive models, namely KNN, SVM, and ANN. The scaler was fitted on the training data only, then applied to the test data to avoid information leakage.

The data were divided using a stratified 80:20 holdout split. The training set contained 120 observations, while the testing set contained 30 observations. Stratification preserved the class proportion in both sets. The testing set contained 9 low-quality samples, 12 medium-quality samples, and 9 high-quality samples. Stratified 5-fold cross-validation was then applied to assess model stability across different partitions. Scores were reported as mean plus or minus standard deviation.

The computational experiment used Python 3.x with NumPy, pandas, scikit-learn, XGBoost, and Matplotlib. NumPy and pandas supported data generation and tabular processing. Scikit-learn

supported preprocessing, train-test splitting, model training, cross-validation, and performance metrics. XGBoost supported the boosted-tree baseline. Matplotlib supported figure generation. The experiment used a standard CPU-based computer and did not require a GPU, camera, lighting unit, or image-acquisition device because the study used numerical simulated data rather than image data.

Random Forest was configured with 200 trees, the Gini criterion, random_state set to 42, balanced class weights, and a minimum of two samples per leaf. Balanced class weights reduced sensitivity to class-count differences. The minimum-leaf setting reduced overly specific splits in the small simulated dataset. Aggressive hyperparameter tuning was not applied because the goal was to build a reproducible baseline workflow rather than an optimized industrial model.

Model performance was compared with K-Nearest Neighbor, Decision Tree, Support Vector Machine, Artificial Neural Network, and XGBoost. KNN represented a distance-based baseline. Decision Tree represented the base learner inside Random Forest. SVM represented a margin-based classifier. ANN represented a nonlinear neural baseline. XGBoost represented a boosted tree-ensemble baseline. This comparison provides a broader evaluation than a comparison with only simple classifiers.

The evaluation used accuracy, weighted precision, weighted recall, weighted F1-score, confusion matrix analysis, and stratified 5-fold cross-validation. Weighted averages were used because the three classes did not have identical sample counts. The formulas were accuracy = (TP + TN) / (TP + TN + FP + FN), precision = TP / (TP + FP), recall = TP / (TP + FN), and F1-score = 2 x (precision x recall) / (precision + recall). Feature importance was analyzed using impurity-based Random Forest importance. Because impurity-based importance can be biased when predictors differ in scale or split opportunities, the interpretation was treated as exploratory. Recent literature recommends validating important variables through complementary methods such as permutation importance, SHAP, or repeated validation before operational use (Rothacher & Strobl, 2024; Saarela & Podgorelec, 2024).

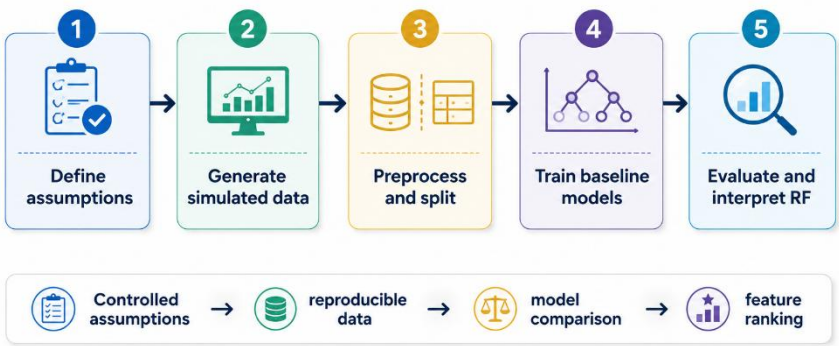

Figure 1. Workflow for silk fabric quality prediction using simulated physical feature data

Figure 1 summarizes the research flow. The process begins with quality assumptions, then moves to simulated data generation, preprocessing, model comparison, and Random Forest interpretation. This flow clarifies that the study simulates the dataset, not the textile production system.

RESULTS AND DISCUSSION

Table 3 shows the holdout and cross-validation results. Random Forest achieved holdout accuracy of 0.87, precision of 0.88, recall of 0.87, and F1-score of 0.87. The 5-fold cross-validation accuracy was 0.85 plus or minus 0.06, while the cross-validation F1-score was 0.85 plus or minus 0.06. These results indicate stable performance in the controlled simulation, although the fold variation shows that the small sample size still affects model stability.

The broader baseline comparison gives a more careful interpretation. Random Forest did not outperform every model on every metric. KNN, SVM, and ANN produced the same holdout accuracy, while SVM and ANN produced slightly higher mean cross-validation accuracy.

Random Forest remained the most suitable primary model for this study because it combined competitive predictive performance, low preprocessing complexity, robustness on small tabular data, and direct feature-importance output. Decision Tree performed lower than Random Forest, which supports the use of an ensemble instead of a single tree. XGBoost produced weaker baseline performance, likely because the small dataset and limited tuning reduced the advantage of boosting.

Table 3. Holdout and stratified 5-fold cross-validation performance

Model	Holdout Acc.	Precision	Recall	F1	5-fold CV Acc.	5-fold CV F1
KNN	0.87	0.87	0.87	0.87	0.84 +/- 0.10	0.84 +/- 0.10
Decision Tree	0.63	0.62	0.63	0.63	0.75 +/- 0.10	0.75 +/- 0.10
SVM	0.87	0.89	0.87	0.87	0.86 +/- 0.07	0.86 +/- 0.08
ANN	0.87	0.88	0.87	0.87	0.86 +/- 0.02	0.86 +/- 0.02
Random Forest	0.87	0.88	0.87	0.87	0.85 +/- 0.06	0.85 +/- 0.06
XGBoost	0.77	0.80	0.77	0.77	0.81 +/- 0.05	0.80 +/- 0.06

Table 3 shows that Random Forest achieved a strong balance between classification performance and interpretability. The results also prevent overclaiming because several models achieved similar holdout scores. The practical advantage of Random Forest lies in its stable ensemble structure and immediate feature-ranking output.

Figure 2 presents the Random Forest confusion matrix. The model correctly classified all nine low-quality test samples. In the medium-quality class, ten samples were correctly classified and two samples were predicted as low quality. In the high-quality class, seven samples were correctly classified and two samples were predicted as medium quality. Errors occurred only between adjacent classes. This pattern is reasonable because simulated medium-quality fabric shares partial feature ranges with both low- and high-quality fabric.

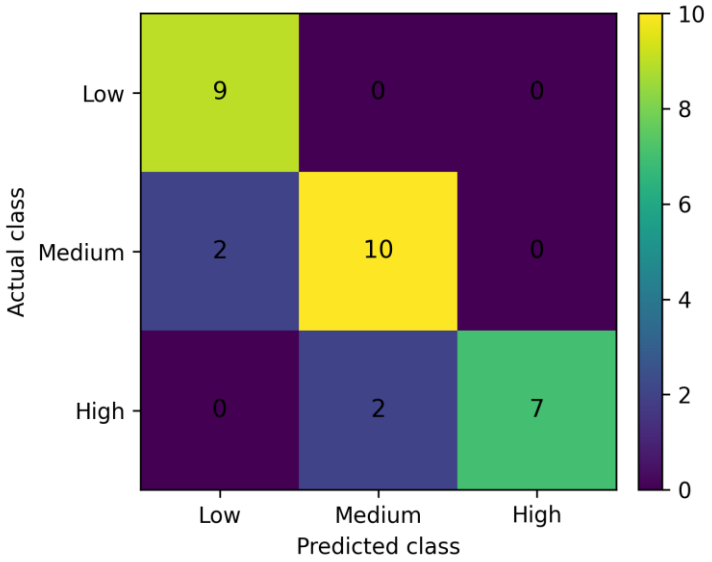

Figure 2. Confusion matrix of Random Forest classification results

Figure 2 confirms that misclassification occurred only between neighboring quality levels. The model did not confuse low-quality samples with high-quality samples. This result supports the assumption that adjacent classes are more difficult to separate when feature ranges overlap.

Table 4 and Figure 3 present the feature-importance results. Surface defects ranked first with an importance value of 0.322. This result follows textile quality-control logic because visible defects directly affect visual acceptance and commercial grading. A fabric with acceptable strength or density can still be downgraded when visible defects appear frequently.

Tensile strength ranked second with an importance value of 0.157. This variable reflects resistance to mechanical force and indicates durability during processing and use. Thread density ranked third with an importance value of 0.156. Denser and more consistent thread structure

supports fabric stability, surface regularity, and structural consistency. These three variables form the dominant predictor group in the simulated model.

Fabric weight, luster, softness, elasticity, and thickness had smaller importance values. These lower values do not mean that the variables are irrelevant in real silk production. They only show that the current simulation assigned stronger discriminative power to surface defects, tensile strength, and thread density. Real production data may produce different rankings because actual quality depends on yarn type, humidity, dyeing, finishing, loom settings, and inspector criteria.

Table 4. Feature importance values generated by Random Forest

Rank	Feature	Importance value
1	Surface defects (count)	0.322
2	Tensile strength (N)	0.157
3	Thread density (thread/inch)	0.156
4	Fabric weight (g/m2)	0.090
5	Luster (1-10)	0.082
6	Softness (1-10)	0.075
7	Elasticity (%)	0.062
8	Fabric thickness (mm)	0.057

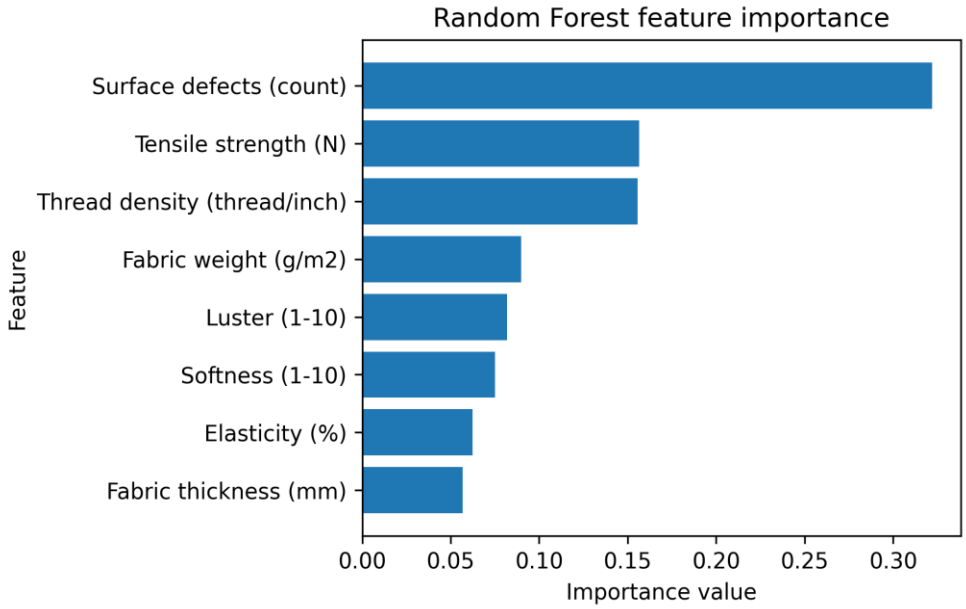

Figure 3. Feature importance of physical material features

Figure 3 visualizes the dominance of surface defects, tensile strength, and thread density. This ranking helps translate the model output into quality-control priorities. In a real setting, these variables should be validated with actual inspection and laboratory data before being used as operational control points.

Compared with recent image-based textile inspection studies, this research takes a different route. Computer-vision models are suitable when the main objective is to localize visible defects in real time (Cheung & Yang, 2024; Lin et al., 2024). This study uses physical material features that can be measured through laboratory or routine material testing. The approach may support early-stage quality-control systems that lack imaging infrastructure. Its value lies in transparent workflow design and variable prioritization, not in replacing visual inspection.

The main limitation is the simulated dataset. The numerical performance cannot be generalized to real silk fabric production. Real data may contain larger measurement noise, class imbalance, inconsistent grading, batch effects, environmental variation, process variation, and disagreement among inspectors. Future research should collect real silk fabric data, report actual measurement protocols, expand the sample size, apply repeated stratified validation, and compare impurity-based feature importance with permutation importance or SHAP. External validation is required before the model can support industrial quality-control decisions.

CONCLUSION

This study developed a Random Forest workflow for predicting silk fabric quality from simulated physical material feature data. The model showed competitive classification performance and produced interpretable feature rankings. Surface defects, tensile strength, and thread density emerged as the dominant predictors because they represent visual acceptance, mechanical durability, and structural consistency. The main contribution is a reproducible simulation framework that links physical-feature assumptions, model comparison, and feature-importance interpretation for early-stage quality-control research. Because the dataset was simulated, the findings cannot be treated as industrial performance evidence. Future studies should validate the workflow with real laboratory or production data, larger samples, additional process variables, and complementary explainability methods.

ACKNOWLEDGEMENTS

The authors thank all parties who supported the preparation of this manuscript. This study was prepared as a methodological demonstration using a controlled simulated dataset and did not receive external research funding.

REFERENCES

ASTM International. (2024). ASTM D5035-11(2024), Standard test method for breaking force and elongation of textile fabrics (strip method). ASTM International.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32. https://doi.org/10.1023/A:1010933404324

Carrilho, R., Peixoto, H., Machado, J., & Abelha, A. (2024). Toward automated fabric defect detection: A survey of recent computer vision approaches. Electronics, 13(18), 3728. https://doi.org/10.3390/electronics13183728

Cheung, W. H., & Yang, Q. (2024). Fabric defect detection using AI and machine learning for lean and automated manufacturing of acoustic panels. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 238(12), 1827-1836. https://doi.org/10.1177/09544054231209782

Ferreira, H. M., Carneiro, D. R., Guimaraes, M. A., & Oliveira, F. V. (2024). Supervised and unsupervised techniques in textile quality inspections. Procedia Computer Science, 232, 426-435. https://doi.org/10.1016/j.procs.2024.01.042

Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on typical tabular data? Advances in Neural Information Processing Systems, 35, 507-520.

ISO. (1977). ISO 3801:1977, Textiles - Woven fabrics - Determination of mass per unit length and mass per unit area. International Organization for Standardization.

ISO. (1996). ISO 5084:1996, Textiles - Determination of thickness of textiles and textile products. International Organization for Standardization.

Kahraman, Y., & Durmusoglu, A. (2023). Deep learning-based fabric defect detection: A review. Textile Research Journal, 93(5-6), 1485-1503. https://doi.org/10.1177/00405175221130773

Li, C., Li, J., Li, Y., He, L., Fu, X., & Chen, J. (2021). Fabric defect detection in textile manufacturing: A survey of the state of the art. Security and Communication Networks, 2021, Article 9948808. https://doi.org/10.1155/2021/9948808

Lin, H., Cai, D., Xu, Z., Wu, J., Sun, L., & Jia, H. (2024). Fabric4show: Real-time vision system for fabric defect detection and post-processing. Visual Intelligence, 2, Article 13. https://doi.org/10.1007/s44267-024-00047-w

Rothacher, Y., & Strobl, C. (2024). Identifying informative predictor variables with random forests. Journal of Educational and Behavioral Statistics, 49(4), 595-629. https://doi.org/10.3102/10769986231193327

Saarela, M., & Podgorelec, V. (2024). Recent applications of explainable AI (XAI): A systematic literature review. Applied Sciences, 14(19), 8884. https://doi.org/10.3390/app14198884

Song, S., Jing, J., Huang, Y., & Shi, M. (2021). EfficientDet for fabric defect detection based on edge computing. Journal of Engineered Fibers and Fabrics, 16, 1-12. https://doi.org/10.1177/15589250211008346

Tu, Y. F., Kwan, M. Y., & Yick, K. L. (2024). A systematic review of AI-driven prediction of fabric properties and handfeel. Materials, 17(20), 5009. https://doi.org/10.3390/ma17205009