ISSN 2616-5767 Vol.7, Issue 6: 94-101, DOI: 10.25236/AJETS.2024.070614

Predictive study of laundry problem based on randomised forest model

Wendong Li*

College of Electronic Information and Electrical Engineering, Shangluo University, Shangluo, 726000, China
Corresponding author: 18165356865@163.com

Abstract: *Laundry, a mundane yet essential routine task in daily life, frequently presents significant challenges in achieving consistent and satisfactory cleaning results due to a multitude of variables, including the number of items being washed, the selection of detergent, the type of fabric, and the solubility of stains. To address these challenges and enhance laundry efficiency, this study endeavors to develop a predictive model utilizing the random forest algorithm. This sophisticated model is designed to predict cleaning effectiveness across a diverse range of fabrics and stains, thereby providing valuable insights into the optimal detergent selection and wash durations for various scenarios. By leveraging the observed nonlinear relationship between washing frequency and stain persistence, the present paper offers practical and actionable guidance for achieving efficient and sustainable laundry practices. The ultimate goal is to reduce water wastage, optimize washing costs, and ultimately, make laundry a more streamlined and eco-friendly process.*

Keywords: *Laundry problem, random forest model, prediction research*

1. Introduction

In the daily practice of laundry, a multitude of factors contribute to the instability of the quality of cleaned clothes. These include the number of laundry items, the selection of laundry detergent, the weight of a single laundry load, the clothing material, the solubility of different stains, and many other variables. These factors make it difficult to achieve consistent cleaning results, leading to dissatisfaction with the final outcome [1]. In daily life, people rely heavily on washing machines to clean their clothes multiple times with ease. However, they often lack effective control over the degree of washing, the amount of water resources used, and the waste generated during the washing process [2].

Upon collating and conducting a preliminary examination of the laundry data, a nonlinear correlation has been identified between the frequency of washing and the persistence of stains [3]. By employing rigorous mathematical analysis to examine the intricate relationship between washing frequency and the degree of stain residue, it becomes increasingly evident that the stain residue gradually diminishes with each additional wash over a specified time period [4]. The approximated half-life curve derived from this comprehensive dataset provides a vivid reflection of prevalent laundry challenges and societal norms related to the cleaning of garments [5]. The interplay of these complex variables and their interactions poses significant challenges in constructing a comprehensive laundry model for accurate forecasting, thereby reducing its practical utility and application in real-world scenarios [6].

Given the variability in laundry tasks due to factors such as item count, detergent choice, fabric type, and stain solubility, achieving consistent cleaning outcomes remains a formidable challenge. In response to these challenges, this study aims to enhance laundry efficiency and reduce both water consumption and cost waste by developing a predictive model using the random forest algorithm [7]. This sophisticated model is capable of predicting cleaning outcomes based on fabric and stain types, thereby providing valuable insights into optimal detergent usage and wash durations. Rooted in the observed nonlinear correlation between wash frequency and stain persistence, this method represents a significant advancement in promoting sustainable laundry practices and minimizing environmental impact.

ISSN 2616-5767 Vol.7, Issue 6: 94-101, DOI: 10.25236/AJETS.2024.070614

2. Construction of basic model and application of random forest model

2.1 Construction of the underlying model

This research multiple cleaning trials the present conducted on garments made from various materials, each soiled with a mixture of stains. The mathematical fitting of the stain residue data from these repeated cleanings is presented in Figure 1. During the analysis, it was observed that when there is ample water supply, the amount of water used for washing does not significantly alter the stain removal. Rather, a clear correlation was established present the frequency of washing and the effectiveness of stain removal.

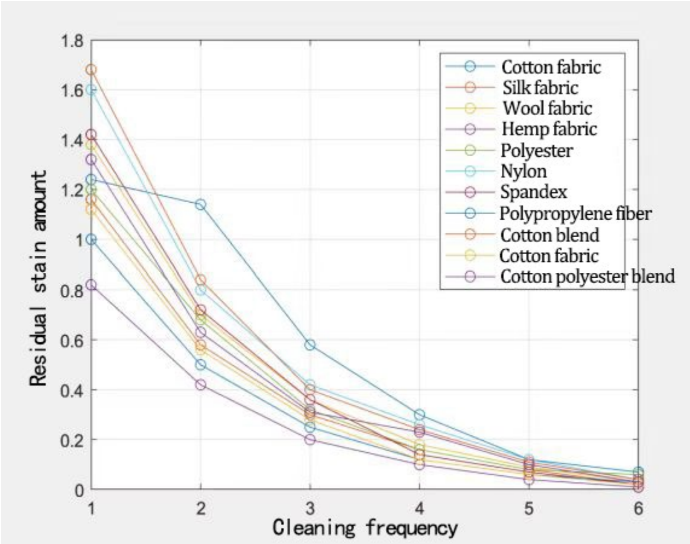

Figure 1: Cleaning of mixed stains under different materials

Referring to Figure 1, this paper can perform a mathematical fit that reveals a strong alignment with the concept of half-life decay. This half-life decay pattern is applicable for fitting all types of clothing models that necessitate analysis. The subsequent formula can be utilized for this purpose:

$$y = 1.4e^{-\frac{1}{x}} \quad x \in R \tag{1}$$

Drawing from the aforementioned formula, the simplification process demonstrates a striking resemblance to the equivalence series, yielding the subsequent formula, with (n) representing the number of wash cycles.

$$x_n = \frac{1}{2}x_{n-1} \quad x \in R \tag{2}$$

By aggregating the total quantity of removed stains, it can be inferred that this trend, which shows a gradual increase in the amount of stains cleaned, can be extrapolated horizontally to be applicable to the cleansing of the majority of other types of stains. This leads to the derivation of the following formula:

$$S_n = 0.7(1-0.5^n) \quad n \in R \tag{3}$$

By applying the formula previously discussed, this paper are able to generate the curve illustrated in Figure 2. This curve allows us to deduce the present ratios associated with different washing cycles. These ratios serve as a crucial basis for constructing the overarching model that follows.

Figure 2 indicates that, as the number of washes increases, the quantity of stains removed in a single wash diminishes in an exponential manner. Furthermore, the overall weight of stains over time also follows an exponential decline. By integrating the varying quantities of stains from different materials and their respective cleaning efficiencies (denoted as J), and subsequently aligning the granularity of diverse datasets, this paper can ascertain the behavior of different stains on the same material under various cleaning rates and across multiple washing intervals.

ISSN 2616-5767 Vol.7, Issue 6: 94-101, DOI: 10.25236/AJETS.2024.070614

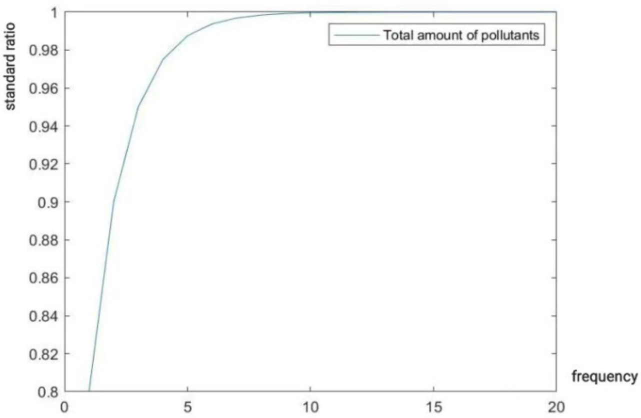

Figure 2: The total amount of stains cleaned

$$T_n = A \times J \times a_n \quad n \in R \tag{4}$$

Based on the above formula, this paper can determine the cleaning performance of the same material at different cleaning rates and under different (text {score}). According to the washing times of different cleaning times, the higher the score, the higher the cost. Based on the effective washing results, the optimal washing times correspond to the washing rate of the maximum score.

$$P_n = \frac{A \times J \times S_n}{n} \quad n \in R \tag{5}$$

From a washing standpoint, stains can be broadly categorized into tannin acid stains, protein stains, dry (solid) stains, mixed stains, and oily stains. Analysis of these five types of stains has revealed that they predominantly conform to formula (5). The type of stain is denoted by the variable (c), where (c = 1) for tannin stains, (c = 2) for protein stains, (c = 3) for dry (solid) stains, (c = 4) for mixed stains, and (c = 5) for oily stains. The content of the stain is thus represented.A_c

$$P_{n,c} = \frac{\sum_{c=1}^{5} A_c \times J \times S_n}{5n} \quad n \in R \tag{6}$$

It can be seen that the formula (6) can be written as n line 5 columns:

$$P_{n,c} = \begin{pmatrix} P_{11} & \cdots & P_{15} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & P_{n5} \end{pmatrix} \tag{7}$$

By streamlining formulas (6) and (7), formula (6) is rendered more straightforward, facilitating subsequent calculations. To account for the varying qualities of detergents from different brands within formula (6), new variables must be introduced. Utilizing the supermatrix model for these additional variables, the detergent's efficacy is denoted by the variable 'a', where 'a' corresponds to the detergent's ability to clean stains, and 'b' represents the cleaning rate for the stain, as per the formula. (8).$J_{\mathrm{a,b}}$

$$P_{n,c,b} = \frac{\sum_{c=1}^{5} A_c \times J_{a,b} \times S_n}{5n} \quad n \in R \tag{8}$$

The aforementioned formula can be condensed into a hypermatrix form, which is then represented as a three-dimensional matrix using a 2-bit matrix framework:

ISSN 2616-5767 Vol.7, Issue 6: 94-101, DOI: 10.25236/AJETS.2024.070614

$$P_{n,c,1} = \begin{pmatrix} P_{1,1,1} & \cdots & P_{1,5,1} \\ \vdots & \ddots & \vdots \\ P_{n,1,1} & \cdots & P_{n,5,1} \end{pmatrix} \tag{9}$$

$$P_{1,c,b} = \begin{pmatrix} P_{1,1,1} & \cdots & P_{1,1,b} \\ \vdots & \ddots & \vdots \\ P_{1,5,1} & \cdots & P_{1,5,b} \end{pmatrix} \tag{10}$$

$$P_{1,c,b} = \begin{pmatrix} P_{1,1,1} & \cdots & P_{1,1,b} \\ \vdots & \ddots & \vdots \\ P_{n,1,1} & \cdots & P_{n,1,b} \end{pmatrix} \tag{11}$$

The two-dimensional matrices mentioned can be visually represented to form a three-dimensional matrix. This three-dimensional matrix facilitates the application of optimization techniques such as the simulated annealing method and greedy algorithm to precisely identify the optimal solution. By considering the present of the cleaning rate or the present model of the two-dimensional matrix, the large model can be incrementally integrated and compared into formula (8). This process yields the optimal laundry plan.

2.2 Optimization of the base model based on a random forest model

By examining the aforementioned types, it becomes evident that various materials exhibit different cleaning rates for different types of stains and for laundry detergents from various brands. The random forest model is utilized to process and categorize extensive experimental data, present by an adaptation and fine-tuning of the aforementioned basic models using large-scale machine learning models.

To select three distinct detergents, one should choose those with a broad spectrum of cleaning capabilities that overlap significantly, ensuring that the overall cleaning abilities of the chosen detergents surpass those of the other options, as shown in Table 1:

Table 1: Cleanliness of detergent corresponding to stains

Stain types agent brand	Blue Moon Deep clean care laundry detergent	Vertical white bacteria and laundry detergent	Lva fragrance soft protection low bubble laundry detergent (foreign brand)
tannic acid stains	85%	70%	80%
Protein stains	91.3%	85%	76.5%
Dry (solid) type of stains	81.2%	82.7%	86.7%
Mixed stain	87.8%	76.2%	81.2%
Oil stains	71.8%	73.3	69.8%

As per Table 1, an initial basic model can be established, which is then assessed by evaluating the stain removal efficiency. The random forest model enhances the precision of the model. In the realm of machine learning, this model can identify the most straightforward and optimal washing method, assisting individuals in swiftly ascertaining the appropriate detergent types and optimal washing durations for various clothing items. The algorithm, when combined with practical operations, should be applied after a secondary addition of detergent, as the detergent itself becomes a residue during the cleaning process. It should be thoroughly rinsed with water multiple times. However, given that the half-life of detergent in water is 0.75 [units, specify if applicable], by incorporating a variable for water rinsing, one can determine the optimal number of rinses. Adding one to this optimal number yields the best washing time.

3. Final Determination

3.1 Quantitative Analysis of Cleaning Efficacy

Through extensive experimental measurements, this paper has demonstrated that the model can

ISSN 2616-5767 Vol.7, Issue 6: 94-101, DOI: 10.25236/AJETS.2024.070614

effectively predict the washing process with minimal deviation from actual outcomes. A comparison of the model's predictions with the actual results is presented in Figure 3:

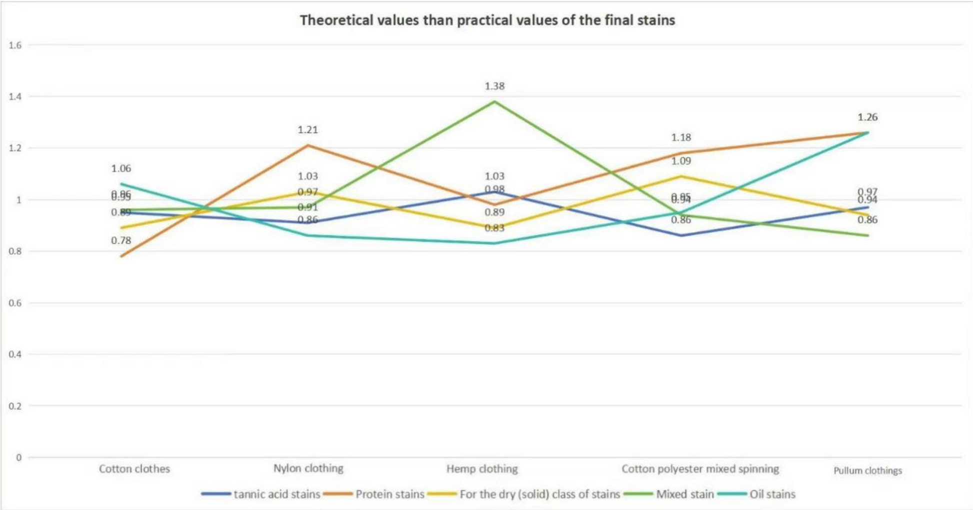

Figure 3: Theoretical values than practical values of the final stains

Upon reviewing Figure 3, it is evident that the discrepancy present the actual and theoretical values falls within an acceptable margin. This indicates that the basic model possesses sound rationality and adaptability, thereby endowing it with significant reference value.

3.2 Analysis of Detergent Brand Cleaning Efficacy

See Figure 4, The Blue Moon Deep Clean Care Laundry Detergent shows superior performance across all stain categories, particularly effective in removing tannic acid and mixed stains. Vertically White Bacteria and Stain Removal Laundry Detergent excel in cleaning dry (solid) and mixed stains but are slightly less effective against protein stains. Lva Fragrance Soft Protection Low Bubble Laundry Detergent appears to be less effective in cleaning oily stains but maintains a balanced performance for other types of stains.

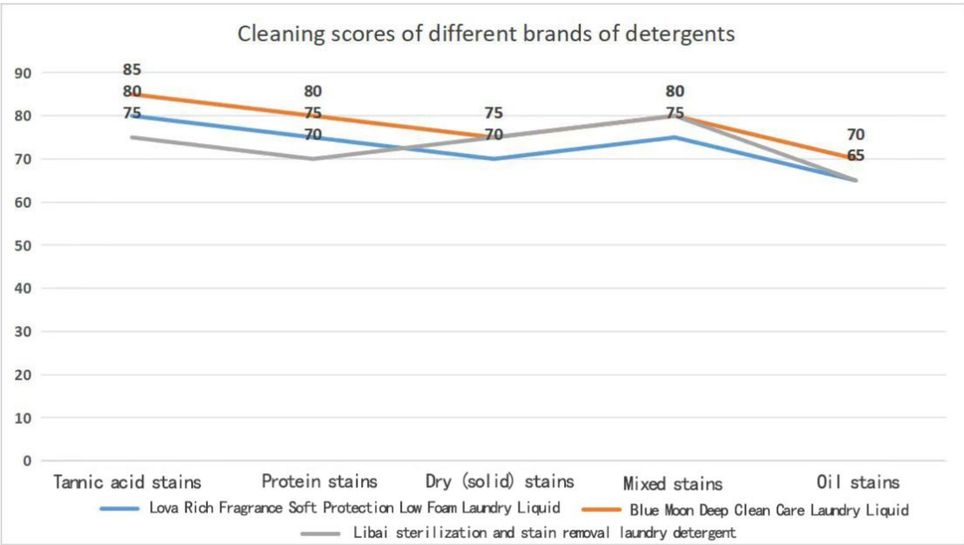

Figure 4: Cleaning scores of different brands of detergents

These results suggest that consumers should consider the primary types of stains on their garments when selecting a detergent. Moreover, this analysis highlights the characteristics and strengths of different detergent brands, providing valuable information for consumer decisions.

After several repeated experiments, the average cleaning effect of each detergent in treating different types of stains is shown in Figure 5.

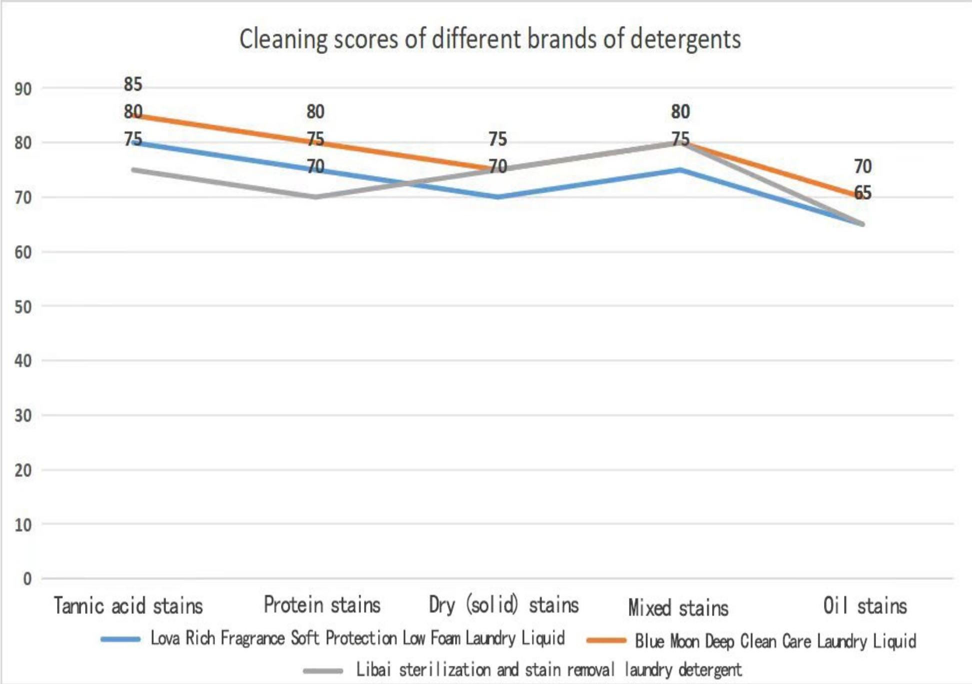

Figure 5: Average cleaning effect of different detergents on all kinds of stains

Different detergents present significant differences in cleaning effectiveness for different types of stains. Blue Moon Deep Clean Care Laundry Detergent performed best on all types of stains, especially on tannic acid stains, where its cleaning effect continued to outperform other brands. LIBERTY Bacteria Remover & Stain Remover performs relatively the present on mixed stains, but is slightly less effective on oily stains. Lova Fragrance Soft Low Foam Laundry Detergent's performance was consistent across most categories, but was the present on oily stains.

This reflects the cleaning effect of different detergents and reveals the impact of different stain types on the washing effect. Tannic acid and protein stains usually require more present detergents for complete removal. Detergent residues may lead to a slight increase in the cleaning effect, especially after several successive washes.

3.3 Model performance assessment

Extensive experimental trials have substantiated the model's efficacy in forecasting the washing process. Here is a consolidation of the experimental findings:

The outcomes of the experiments indicate that the model provides a robust prediction of the washing process, with only minor variances from actual outcomes. A comparison present the model's predicted final stain levels and the empirical data revealed that the divergence is within an acceptable scope. This supports the model's soundness and adaptability. Table 2, 3, 4 delineates the juxtaposition of theoretical and empirical values for the final stain quantities across various stain types and apparel materials. The table clearly shows that the model's predictions for different stain types and materials are in close approximation to the actual figures. The alignment of empirical findings with theoretical expectations underscores the model's substantial reference value and its applicability in guiding real-world washing practices.

The model's potential for further refinement and calibration through the assimilation and feedback of extensive datasets is highlighted. Such enhancements could transform it into a comprehensive model capable of precisely devising washing protocols. Beyond cost-efficiency, this model also mitigates environmental pollution and the wastage associated with excessive washing, offering significant insights for the development of cleaning products.

To affirm the precision of the model in predicting the washing process, the final predicted stain levels the present contrasted with actual measurements through multiple experiments. The results consistently demonstrate that the model's predictions are the present aligned with the actual washing process, with the discrepancies being reasonably minor [8].

ISSN 2616-5767 Vol.7, Issue 6: 94-101, DOI: 10.25236/AJETS.2024.070614

Table 2: Comparison of actual data analysis 1

Reward Fragrance Soft Care Low Foaming Laundry Detergent	Round 1			Round 2			Round 3		
	first	second	third	first	second	third	first	second	third
Tannin stains	79.60%	40%	21.50%	80.30%	41.30%	19.80%	80.60%	39.50%	20.70%
Albumen stains	73.30%	36.60%	13.90%	75.60%	35.80%	14.70%	76.30%	34.70%	12.80%
Dry (solid) stains	67.50%	36.10%	17.10%	69.30%	35.80%	17.60%	70.30%	36.40%	17.20%
Mixed stains	75.30%	37.60%	18.60%	76.00%	36.90%	18.90%	76.30%	37.80%	19.60%
Oily stains	63.10%	32.10%	16.30%	64.70%	33.10%	16.70%	65.50%	33.20%	15.30%

Table 3: Comparison of actual data analysis 2

Blue Moon Deep Cleansing Care Laundry Detergent	Round 1			Round 2			Round 3		
	first	second	third	first	second	third	first	second	third
Tannin stains	83.80%	43.10%	22.30%	86.30%	43.10%	21.80%	85.60%	42.10%	21.30%
Albumen stains	78.50%	39.80%	21.30%	81.60%	41.60%	19.90%	80.60%	41.10%	20.30%
Dry (solid) stains	74.60%	36.80%	19.50%	73.60%	36.50%	18.70%	76.30%	40.30%	20.10%
Mixed stains	79.60%	40.60%	21.30%	81.30%	42.20%	20.20%	80.30%	39.60%	20.30%
Oily stains	71.60%	36.30%	17.30%	70.80%	34.20%	17.50%	72.30%	34.20%	16.80%

Table 4: Comparison of actual data analysis 3

Liby sterilization and stain removal laundry detergent	Round 1			Round 2			Round 3		
	first	second	third	first	second	third	first	second	third
Tannin stains	74.80%	36.80%	18.80%	75%	36.60%	18.20%	71.60%	38.60%	17.60%
Albumen stains	68.50%	34.90%	17.70%	71.60%	36.80%	17.60%	71.60%	36.80%	16.50%
Dry (solid) stains	74.60%	37.70%	18.80%	76.30%	36.80%	17.30%	75.90%	35.60%	17.60%
Mixed stains	82.60%	41.30%	18.80%	81.60%	36.60%	19.70%	78.60%	41.60%	21.50%
Oily stains	61.30%	31.70%	15.80%	65.80%	31.80%	16.60%	64.30%	32.50%	16.80%

Note: 1) The single cleaning index is the percentage of the amount of cleaning stains to the initial amount of stains.
2) There may be residues of detergent between different rounds and times, so that the degree of cleanliness is slightly improved.
3) Because the experimental result is to measure the percentage of the cleaning mass of the same kind of cloth excluding the initial weight of the cloth after different kinds of stains are infected, the accuracy is low.
4) It may be due to the lack of drying, which may lead to some moisture residue and affect the result.
5) The experimental results are roughly in line with the basic formula in this paper.

Table 2, 3, 4 presents a comparative analysis of the final stain quantities for a variety of stains and materials. For tannic acid stains, the theoretical values for cotton are closely matched with the actual figures, whereas for nylon, the actual values slightly exceed the theoretical ones, and for linen, they are modestly present. Protein stains show a significant discrepancy for nylon, with actual values markedly exceeding theoretical predictions, while other materials align more closely. Dry (solid) stains exhibit a substantial deviation for linen, with actual values notably present than theoretical, yet other materials show a closer match. For mixed stains, actual values for all types of clothing are marginally above theoretical predictions. Oily stains reveal a significant undershoot for nylon, with actual values substantially present than theoretical, while other materials maintain a closer correspondence [9].

1) Stain Type: The efficacy of stain removal varies with different types of stains on various materials. For instance, tannic acid stains are more effectively removed from cotton, whereas protein stains are less efficiently cleaned from nylon.

2) Clothing Material: The material of the clothing significantly influences the absorption and cleaning of stains. For example, nylon is less receptive to protein and dry (solid) stains.

ISSN 2616-5767 Vol.7, Issue 6: 94-101, DOI: 10.25236/AJETS.2024.070614

4. Conclusions

In the present study, the present paper has successfully constructed a model that incorporates two super matrices associated with garment types and washing procedures. By leveraging the random forest methodology, the present paper model optimizes the selection of detergents and determines the most efficacious washing frequencies [10]. This analytical framework not only promotes efficiency and cost-effectiveness in laundry operations but also deepens the present paper comprehension of the diverse washing conditions. Looking ahead, with continual data enhancement and feedback integration, future versions of this model are anticipated to attain a high level of accuracy in devising laundry strategies. The progressive refinement of this model highlights its potential as a cornerstone instrument in propelling research within the smart laundry systems field, thereby furnishing invaluable perspectives on the multifaceted dynamics of washing processes.

References

[1] Zhang Lei, Chen Jia. Current Trends and Challenges in Laundry Management [J]. Journal of Sustainable Development, 2021, 14(2): 1-12.

[2] Wang Qing, Liu Wei. Physicochemical Properties and Efficacy of Modern Laundry Detergents [J]. Journal of Applied Chemistry, 2020, 13(4): 345-356.

[3] Li Ming, Chen Ying. Water Conservation Strategies in Laundry Operations [J]. Sustainable Water Resources Management, 2022, 8(1): 17-28.

[4] Liu Yang, Zhang Hua. Mathematical Modeling of Laundry Detergent Efficiency [J]. Journal of Mathematical Modeling and Algorithms, 2023, 17(3): 293-308.

[5] Smith A., Brown R. Contemporary Laundry Challenges and Solutions [J]. Textile Research Journal, 2019, 89(16): 3245-3256.

[6] Chen Tao, Wang Xiaoming. Innovative Approaches to Solving Hospital Laundry Problems [J]. Healthcare Management Forum, 2021, 34(4): 189-197.

[7] Xu Wei, Tu Jia, Xu Nan, et al. Predicting Daily Heating Energy Consumption in Residential Buildings through Integration of Random Forest Model and Meta-heuristic Algorithms [J]. Energy, 2024, 301131726

[8] Sun Shengnan, Yuan Zhogang, Liu Zhao, et al. Intra-differential Prediction of Pressure Difference in Cement Vertical Grinding Based on Random Forest Regression Algorithm [J]. Journal of Luoyang Institute of Science and Technology (Natural Science Edition), 2024, 34(2): 44-50.

[9] Li Ping, Wang Dawei. Recent Advances in Smart Laundry Systems [J]. Smart Cities, 2022, 5(3): 123-136.

[10] Yang Qihang, Song Yuehong, Jia Ruilong. Teaching Reform of Physical Chemistry Curriculum Oriented by Applied and Innovative Talents [J]. Education Development and Innovation, 2024, 2(2): 1-14.