Research Article GU J Sci 37(1): 326-337 (2024) DOI: 10.35378/gujs.1185130

Gazi University

Journal of Science

http://dergipark.gov.tr/gujs

Developed ABCLASS-Miner Classification Algorithm Based Rule Extraction for Denim Fabrics

Gözde KATIRCIOGLU[1], Emel KIZILKAYA AYDOGAN[1,*], Esra AKGUL[2], Yılmaz DELICE[3]

[1] *Department of Industrial Engineering, Faculty of Engineering, Erciyes University, Kayseri, Turkey*
[2] *Department of Industrial Design Engineering, Faculty of Engineering, Erciyes University, Kayseri, Turkey*
[3] *Department of International Trade and Logistics, Faculty of Applied Sciences, Kayseri University, Kayseri, Turkey*

Highlights

- This paper is based on a new classification rule inference algorithm proposed for denim fabrics.
- Aims to figure out the relationship between denim fabric production and quality parameters.
- The categorization accuracy was improved to a very high level of precision.

Article Info

Received: 06 Oct 2022
Accepted: 03 May 2023

Keywords

Data mining
Artificial bee colony
Classification algorithm
Denim fabrics

Abstract

Obtaining and storing large amounts of data have become easier with the rapidly developing information technologies (IT). However, the data generated and collected, which are irrelevant in and of themselves, become useful only when they are analyzed for a specific reason. Data mining may transform raw data into useful information. In the present study, classification and analysis of denim fabric quality characteristics according to denim fabric production parameters were carried out. The present study proposes a new classification rule inference algorithm. The suggested approach is mostly based on Artificial Bee Colony Optimization (ABC), a swarm intelligence meta-heuristic. In each step of the algorithm, there are two phases called the employed bee phase and the onlooker bee phase. This algorithm has been compared with the classification algorithms in the related literature. This proposed algorithm is a new data mining tool that intelligently combines various metaheuristic and neural networks and can generate classification rules. The results indicate that the proposed data mining algorithms may be highly useful in determining weight and width in denim fabric manufacture.

1. INTRODUCTION

Today, information technologies are developing rapidly and facilitate the acquisition and storage of large amounts of data. Processing raw data and discovering useful information derived from the processed data have encompassed many fields such as engineering, medicine, science, public and private sector enterprises. In this context, the opinion of data mining plays a significant role [1]. Data mining is the process of extracting valuable and usable information from existing data [2]. Namely, it is to make predictions about the future and set logical rules based on available data [3]. The purpose of data mining is not to organize the existing data, but to try to offer a different perspective and solution by extracting meaningful analyses from the existing data sets. Inferences in data mining are made by using many techniques such as clustering, classification, association rules. Among these data mining tasks, classification rule extraction has received great attention in recent years.

Data mining is now used in a variety of sectors, including marketing, financial analysis, clinical medicine, finance, and fraud detection. Considerable amount of data such as the raw material of the product, machine settings, quality parameters are produced and stored in textile sector. There is a constant demand from the market to process these data and discover usable information. Data mining used in the textile industry is an interdisciplinary field [4] such as decision support systems [5], recommendation systems [6], visual data analytics [7], database management system [8], field-based data mining [9]. In recent years, many scientists

*Corresponding author, e-mail: ekaydogan@erciyes.edu.tr

have studied rule extraction for classification in the related literature. Xiang et al. [10] investigated the analyses of yarn tenacity to give the properties of the fibers with a rule extraction method based on Rough sets theory (RST). Deng et al. [11] established classification rules and used an artificial neural network approach to classify the pilling degree of knitted, woven, and non-woven materials. Kumar [12] analyzed and processed relevant data to find fabric defects, which were classified by means of computational techniques with images obtained from the video of the rolling knit fabric. Xiang et al. [13] proposed an intelligent control model (ICM) based on Rough sets theory for the spinning process. Mozafary et al. [14] estimated yarn quality by using data mining methods including artificial neural networks. Pang and Chan [15] proposed a data mining-based algorithm to minimize the overall travel distances for both shelving and order picking operations. Mohanty and Bag [16] proposed a method based on GLCM and an Association Rule classifier in order to identify textile defects. Lizarraga-Morales et al. [17] proposed an approach using local binary features and a rule-based classification for the defect detection in textiles with a patterned texture. Revathy et al. [18] proposed a methodology for fabric fault detection and classification of defects. Das and Ghosh [19] generated decision rules using rough set theory to classify neps in cotton yarn. Bhuvaneshwarri et al. [20] applied rough set theory and data mining techniques to estimate fabric width. Bhuvaneshwarri et al. [21] proposed an optimal feature selection based on rough computation methodology for fabric width estimation.

Metaheuristic algorithms have also been employed extensively in the textile sector in recent years. Metaheuristic algorithms search for optimal feature rules in very large search spaces and are powerful algorithms for solving global optimization problems of constant space. They have achieved promising performance in solving the majority of real-world optimization problems. In the literature, metaheuristic algorithms in the textile industry have been applied in many studies such as classification, prediction and defect detection. Deng et al. [22], for example, introduced a novel intelligent multi-criteria optimization technique for the production of multimodal textile materials. Fu et al. [23] used such algorithms to detect weft in glass fiber textile machine with Rule-Based Ant Colony algorithm. Zhao et al. [24] suggested an enhanced Ant Colony Optimization for foreign fiber feature classification in cotton. Xue et al. [25] presented a smart approach utilizing fuzzy set theories and rough set to take primary visual aspects for the explanation of textile tactual properties. Zhao et al. [26] offered three metaheuristic-based techniques to solving the feature selection issue of cotton foreign fibers: Particle Swarm Optimization, Ant Colony Optimization, and Genetic Algorithm. Amor et al. [27] investigated the functional qualities of titanium dioxide nanoparticles on cotton fabric using a novel technique based on the combination of an Artificial Neural Network and the Crow Search Algorithm (MLP-CSA). But nevertheless, metaheuristic algorithms for rule-based categorization of fabric quality characteristics have not been investigated in the literature.

Denim fabric is a coarse woven fabric made of 100% cotton yarn or predominantly cotton with a twill weave style that is popular among people of all ages, including children, women, and men. Denim textiles, which come in a variety of weights, often feature 3/1 or 2/1 warp twill, for example. They are woven using warp threads colored in blue, navy blue, black, and other colors with white weft threads. The typical denim fabric is rigid. The flexibility attribute, on the other hand, may be obtained by employing elastane material in its manufacture. The yarn structure employed in the manufacture of denim fabric and the finishing procedures utilized can result in non-shrinkage.

Figure 1 illustrates a flowchart of the denim fabric production process. The spinning of yarn is the first step in the manufacture of denim. Warping can begin once the yarn has been manufactured. The warp sheets can be dyed using indigo colors after they have been warped. After coloring, the ropes can be dried and readied for the weaver's beam using drying cylinders. Desizing and other preparing operations are used after the weaving process. After all of these procedures, the cloth is ready to be made into garments. Various washing processes can be used [28, 29].

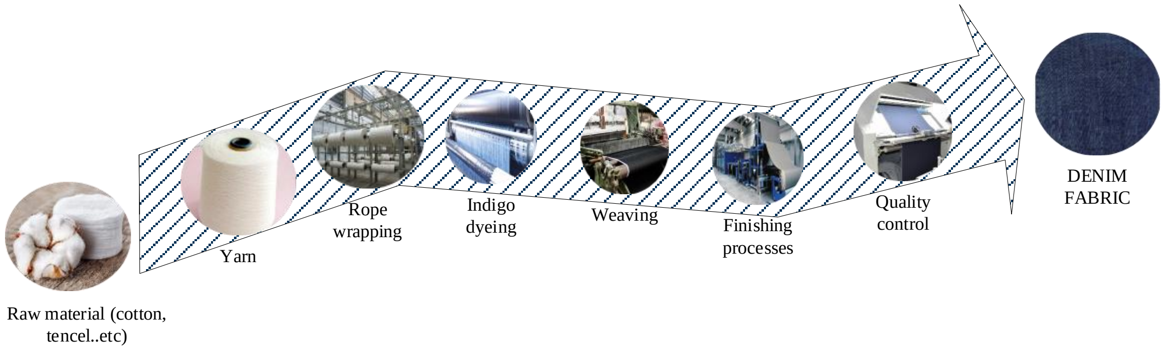

Figure 1. *Process flowchart of denim fabric manufacturing [30]*

In the present study, data mining is based on a new classification rule inference algorithm proposed for denim fabrics. To extract classification rules, the suggested rule extraction approach is primarily rooted in Artificial Bee Colony Based Optimization (ABC). The purpose of this research is to determine the link between denim fabric manufacturing factors and quality attributes. The organization of the present study is as follows: In section 2, the main aspects of data mining are investigated. Next, in section 3, a classification problem is explained. In section 4, the classification algorithms and developed artificial bee colony-based classification algorithms are explained. In section 5, the newly developed ABC based classification algorithm is applied to explore the inferences between denim fabric production parameters and quality properties. Finally, conclusions will be provided.

2. CLASSIFICATION IN DATA MINING AND CLASSIFICATION ALGORITHMS

Classification is a frequently used data mining activity that uses preexisting information to predict the category of future data items. In classification, a training set is often used to train the model, which is subsequently evaluated on a test set. Classification is an example of supervised learning in machine learning language. In the literature, several classification methods have been created. Based on a given input, the classification algorithms attempt to discover correlations between the characteristics of the data in the training set and predict the unknown target attribute value. In classification, a confusion matrix can be used to predict how well an algorithm will perform. Table 1 contains a sample confusion matrix.

Table 1. *The confusion matrix*

		Estimated Class	
		Class +	Class -
Actual Class	Class +	True Positive (TP)	False Negative (FN)
	Class -	False Positive (FP)	True Negative (TN)

In this matrix, TP represents the positive predicted value being positive, TN represents the negative predicted value actually being negative, FP represents the negative predicted being actually positive, FN represents the positive predicting being actually negative. Using these four cases from the confusion matrix, a variety of performance criteria could be calculated. The ratio of correctly categorized items in all data is one of these criteria. Precision is defined as the ratio of total positive elements to properly categorized positive samples. The ratio of properly categorized positive samples to the total number of positive samples is known as recall. The F-measure is the harmonic mean of accuracy and recall criteria. For this analysis, the performance criteria listed in Equations (1)-(4) were applied:

$$\text{Accuracy}=\frac{\text{TP+TN}}{\text{TP+TN+FN+FP}}=\frac{\text{number of correctly classified samples}}{\text{total number of samples}} \tag{1}$$

$$\text{Recall}=\frac{\text{TP}}{\text{TP+ FN}}=\frac{\text{number of correctly classified positive samples}}{\text{total number of positive samples}} \tag{2}$$

$$\text{Precision}=\frac{\text{TP}}{\text{TP+ FP}}=\frac{\text{number of correctly classified positive samples}}{\text{number of positively classified samples}} \tag{3}$$

$$\text{F-measure}=\frac{\text{2*Recall*Precision}}{\text{Recall+Precision}}\ . \tag{4}$$

While evaluating performance of the model, apart from these criteria, the ROC-area value can also be considered. The Roc Area gives the area under the ROC curve. The ROC curve enables the comparison of the efficacy of the tests, the determination of the discriminative power of the tests, and the determination of the appropriate positivity threshold. [31]. The area under the ROC curve determines the accuracy of the test in distinguishing between true and false classification.

In the present study, an ABC based classification algorithm was developed. Performance comparisons of the developed algorithm were made on the production parameters data sets obtained from the aforementioned company with J48, random tree, random forest, Hoeffding tree, naive Bayes, simple logistic, decision table classification algorithms.

Among these algorithms, The J48 is one of the most well-known and widely used decision tree based algorithms. This algorithm uses information to gain the ratio as the test attribute selection criteria. The information gain of each column in the dataset is divided by the information gain of the class to calculate the gain of each column. For each data set, the feature with the highest information gain rate is selected [32]. The random tree algorithm has no pruning and creates a tree for a specific number of randomly selected properties at each node. The generated tree is randomly selected from the possible tree set. Each tree in the tree set has an equal chance of being tried [33]. The random forest developed by Breiman is a supervised learning algorithm and performs well at classifying large amounts of data. This algorithm builds the decision tree by randomizing the divide at each node. It contains many single and unpruned decision trees. The class with the maximum value in the decision forest is picked as the final choice [34]. The Hoeffding tree method is a decision tree classification method that may be used successfully on enormous data sets by analyzing each data set just once. For the building and evaluation of the decision tree, the method employs the Hoeffding bound. Hoeffding limits are used to determine the number of samples to be performed in order to achieve a specific degree of confidence [35]. Naive Bayes tree is a well-known statistical classifier based on Bayes' theorem, in which each feature is conditionally independent and conditionally dependent on each other [36]. It includes probabilistic induction method [37]. The simple logistic is a classification algorithm that creates linear logistic regression models. The goal of a regression analysis is to construct a model predicting the value of another parameter utilizing specific factors. Better results could be obtained with the Boosting process in logistic regression [38]. The decision table is a classification algorithm that creates a set of If-Then rules. A tree structure is created. This algorithm evaluates subsets of features using the best first search and it can use cross validation for evaluation. This algorithm is one of the easiest and simplest classification algorithms to understand [39].

3. DEVELOPED ABCLASS-MINER CLASSIFICATION ALGORITHM

In the proposed study, an ABC based algorithm has been developed so as to conduct a classification process on a data set consisting of quality attributes. Thus, a system that can achieve effective results and can be integrated into any decision support system has been developed. The proposed algorithm could be operated at any time, giving the opportunity to obtain results again by taking into account the current market data. Immediate quality values will be useful for decision-makers in the production department.

In the initial step of the proposed algorithm, all variables are determined, and related initial values are assigned. Since the provided algorithm is a classification algorithm, each class in the structure is considered separately. All the rules for each class are determined by running the whole algorithm sequentially for each class. The proposed algorithm repeats for each class for a certain predetermined number of steps. The data to be analyzed by the algorithm is subjected to various preprocessing and discretization processes.

Each step of the algorithm includes three stages, the employed, the onlooker, and the scout bee stages. In the algorithm that starts with the employed bee stage, random values are initially generated. In each stage the best solution is kept. After the alternative solutions are obtained at the initial stage, the classification process is applied to the training part of the input data. In each solution of the algorithm, there are 3 sub-items of information: operator, value, and type defined for each input parameter. The general structure of the ABC solutions used to derive the classification rules and a rule representation are given in Table 2.

In the first line of the attribute values of a sample solution, values indicating whether that attribute is passive (0) or active (1) are kept. Of these attributes, only the active ones are taken into account in the classification process. In the second line of the rule structure of the example solution, one of the comparison operators <=, > or = is randomly assigned depending on the type of that attribute (binary, categorical, continuous). The third line of the rule structure of the example solution stores the value ranges that the actively determined attributes will have. These ranges of values are randomly generated between the lower and upper limits that each attribute can take. In the fourth line of the rule structure of the sample solution, there is a randomly assigned value indicating which individual rule in 'and' or 'or' groups specified by the active attribute will take place. With this structure, rules with research diversity are produced.

Table 2. *Structure of an ABC solution and the representation of a rule*

		Attributes					
		X_1	X_2	X_3	X_4	X_5	X_6
1	Active/passive	0 or 1					
2	Operator	= , <= or >					
3	Value	values in certain ranges					
4	Type	and / or					

The basic steps of the proposed ABCLASS-miner classification algorithm are given in Table 3.

Table 3. *Developed ABCLASS-miner classification algorithm steps*

Step 1. Define all variables and arrays
For Apply for each class in the problem data
Step 2. Start the employed bees with random values within the allowed range
While (In case the maximum number of steps is not reached) do
Step 3. Employed bee stage
While (In case the maximum number of worker bees is not reached) do
Determine parameter values of candidate solution X_i
Apply the classification procedure to the training set of the problem data
Keep the best solution obtained
End while
Step 4. Generate Onlooker Bees with swap / insert operators
Step 5. Onlooker bee stage
While (In case the maximum number of onlooker bees is not reached) do
Determine parameter values of candidate solution X_i
Apply the classification procedure to the training set of the problem data
Keep the best solution obtained
End while
Step 6. Generate Employed Bees with crossover / mutation operators
Step 7. Perform the scout bee procedure
Generate worker bees with random values whose value does not improve by the number of scout limit steps
End while
End For
Step 8. Combine the rules obtained for all classes
Step 9. Apply the rules obtained to the test set of the problem and write the solution

The parameter values determined intuitively and used in the algorithm are given in Table 4. Using 10-fold cross validation, the data was separated into training and testing data arrays. The training set is employed to learn the algorithm, while the test set is used to evaluate the predictive power of the rule set acquired from the proposed algorithm. The dataset is arbitrarily subdivided into 10 data subsets, with each partition being run once. Each execution of a distinct partition is utilized as the testing set, and the remaining dataset is combined to form the training set [40].

***Table 4.** The parameters used in proposed ANN based ABC algorithm*

Parameters	Value
Number of bee	1000
Number of iteration	1000
Worker/Onlooker rate	0.5
Scout limit	25
Crossover rate	0.8
Mutation rate	0.2

4. THE CLASSIFICATION RULE EXTRACTION WITH ABCLASS-MINER ALGORITHM FOR DENIM FABRICS

4.1. Data Set

The data consists of 330 rigid denim fabrics. The warp and weft yarn number, warp and weft yarn type, weft density and weaving type parameters, which are the production parameters of denim fabric, have been determined as input data. Basically, the study was focused on different yarn types such as cotton, elastane, PES, polyester, tencel with different warp and weft number. Also, these denim fabrics were produced with three weaving types as 2/1 twill, 3/1 twill (3/1s / 3/1z), and broken twill. Weft density data was selected in the range of 13-27 per/cm. Warp and weft yarn number data was selected in the range of 6-30 ne.

The two quality parameters that most affect the sales of fabrics, namely fabric weight and fabric width, are considered as output.

4.2. Preprocessing of Denim Fabric Data

The first step is to combine and clean the data set for the algorithm to give the correct result. The data set was arranged by removing the missing information in the data set. By combining the data, meaningful data were collected in the same file for analysis. Achieving efficient results in the data cleaning and transformation of attributes reduction phases increases the quality of the analysis. Increasing the reliability of the result depends on the classification of the data and the correct categorization of the variables when evaluating them as dependent or independent. The general properties of the determined denim fabrics can be generalized as warp thread number, weft thread number, raw material, weft density and weaving type. The physical and mechanical test data of the examined denim fabrics were obtained by the company. In order to select data mining algorithms in accordance with the data set and to reach useful information, the existing variables must be interpreted correctly.

4.3. The Research Findings and Discussion

The weight is among the significant factors to determine the denim fabric prices. The weight of denim fabric is divided into 3 classes as low weight, medium weight and high weight in the market. The weight can be affected by many factors that form the fabric structure such as fiber type, constructure, density, yarn number, washing type, etc. [41]. The weight per unit values of the denim fabrics have been measured before washing and after 3 household washing processes according to the standards, respectively. The rules between weight, after washing weight and production parameters results have been presented in Table 5 and Table 7. The performance metric values were obtained with J48, random tree, random forest, Hoeffding tree, naive Bayes, simple logistic, decision table and the developed ABCLASS-miner classification algorithms. The performance metric values obtained for the product weight are shown in Table 5.

***Table 5.** The obtained results of classification algorithms for product weight*

	Accuracy	Precision	Recall	F Measure	Roc Area
Random tree	89.00%	0.89	0.89	0.89	0.97
J48	85.05%	0.85	0.85	0.85	0.95
Simple lojistik	79.00%	0.79	0.79	0.79	0.90

Decision table	74.91%	0.76	0.75	0.75	0.88
Naive Bayes	74.39%	0.76	0.75	0.74	0.87
Hoeffding tree	70.61%	0.71	0.71	0.71	0.83
Suggested algorithm	94%	0.94	0.94	0.94	0.98

Table 5 shows that the suggested method achieves the best results (94%) in terms of accuracy, precision, sensitivity, F criteria, and Roc Area performance metrics. For real-world issues, the achieved accuracy is fairly high. The random tree algorithm provides a result of 89% and it is followed by J48, simple logistic, decision table and naive Bayes, respectively. The Hoeffding tree algorithm provides a result of 70.61%. The Hoeffding tree algorithm gave the lowest result. The extracted best rules for the weight of the denim fabric by the proposed algorithm are illustrated in Table 6.

Table 6. *Examples of extracted rules with production parameters for product weight*

Rule 1	IF (the weft yarn is 100% cotton) and (7.1< the weft yarn number<10.35) and (7.05< the warp yarn number<10.15) and (the weaving type is 3/1 Z) and (the weft density >= 14.85), THEN the weight of the product is high.
Rule 2	IF (the weft yarn is 100% cotton) and the weaving type 2/1 Z) and (the weft yarn number < 7.75), THEN the weight of the product is high.
Rule 3	IF (the weft yarn is 100% cotton) and (the weft yarn number < 10.35) and (the weaving type is Broken Twill), THEN the weight of the product is high.
Rule 4	IF (the weft yarn is 100% cotton) and (the weft yarn number< 10.35) and (the weave type is 3/1 S), THEN the weight of the product is high.
Rule 5	IF (the weft yarn is a mixture of 95% cotton and 5% elastane) and (the weft yarn number <13), THEN the weight of the product is high.
Rule 6	IF (the weft yarn is a mixture of 96% cotton and 4% elastane) and (the weft yarn number<11) and (the weft density>=19.15), THEN then the weight of the product is high.
Rule 7	IF (the weft yarn is a mixture of 97% cotton and 3% elastane) and (11< the weft yarn number<13) and (the warp yarn number<8), THEN the weight of the product is high.
Rule 8	IF (the weft yarn is 100% cotton) and (7.1< the weft yarn number<10.35) and (the weaving type is 3/1 Z) and (7.05< the warp yarn number<10.15) and (the weft density< 14.85), THEN the weight of the product is medium.
Rule 9	IF (the weave type is 3/1 Z) and (the weft yarn number< 7.1) and (the weft density< 16.5) and (the warp yarn number>= 10.15), THEN the weight of the product is medium.
Rule 10	IF (the weft yarn is 100% cotton) and (the weaving type is 2/1 Z) and (7.75< the weft yarn number< 10.35), THEN the weight of the product is medium.
Rule 11	IF (the weft yarn is 100% cotton) and (10.3<= the weft yarn number<17.5) and (the warp yarn is 100% cotton) and (8.75< the warp yarn number<10.15), THEN the weight of the product is medium.
Rule 12	IF (the weft yarn is a mixture of 95% cotton and 5% elastane) and (the weft yarn number>=13), THEN the weight of the product is medium.
Rule 13	IF (the weft yarn is a mixture of 96% cotton and 4% elastane) and (the weft yarn number<11) and (18.25<= the weft density<19.15) and (the warp yarn number< 8.35), THEN the weight of the product is medium.
Rule 14	IF (The weft yarn is 100% polyester) and (the weft density>=19.15) and (the warp yarn number< 12.15), THEN the weight of the product is medium.
Rule 15	IF (the weft yarn is 100% cotton) and (10.3<= the weft yarn number<17.5) and (the warp yarn is 100% cotton) and (the warp yarn number>=10.4) and (the weft density<20.15), THEN the weight of the product is low.
Rule 16	IF (the weft yarn is a mixture of 96% cotton and 4% elastane) and (the weft yarn number >= 11) and (the warp yarn number>= 12.5), THEN the weight of the product is low.

According to the rules, it has been observed that as the warp and weft yarn number increases, the fabric weight decreases; however, weaving type and fiber type change are also effective on fabric weight. As the weft density increases, the fabric weight increases.

In Table 7, the classification performance comparison results of the algorithms used for home type washing weight are given.

Table 7. *Results of classification algorithms for home type washing weight*

	Accuracy	Precision	Recall	F Measure	Roc Area
J48	79.88%	0.79	0.79	0.79	0.77
Random tree	77.00%	0.77	0.77	0.77	0.78
Decision table	76.52%	0.76	0.76	0.73	0.65
Simple lojistik	74.39%	0.73	0,74	0.71	0.67
Naive Bayes	69.81%	0.71	0.69	0.68	0.68
Hoeffding tree	69.81%	0.71	0.69	0.68	0.68
Suggested algorithm	93.67%	0.93	0.93	0.93	0.90

Table 7 shows that the proposed algorithm yields the maximum result (93%) according to the accuracy, precision, sensitivity, F criterion performance metrics. J48 algorithm yielded a result of 79%, and it is followed by random tree, decision table, simple logistic and naive Bayes, respectively. The Hoeffding tree algorithm provides a result of 69.81%. The extracted best rules for the weight of the denim fabric by the proposed algorithm are illustrated in Table 8.

Table 8. *Examples of extracted rules with production parameters for home type washing weight*

Rule 1	IF (9.5 < the warp yarn number< 10.2) and (the warp yarn material 100% cotton) and (the weft yarn material 88% combed, 9% elastane, 3% PES or 73% cotton, 17% PES, 10% elastane) and (14.8< the weft density <= 21.5) and (the weave type is 3/1 or 3/1 Z), THEN the weight of the washing is high.
Rule 2	IF (the weft density> 21.5) and (7.4< the warp yarn number<= 9.5) and (the weaving type 2/1 Z) and (the weft yarn number<= 17), THEN the weight of the washing is medium.
Rule 3	IF (15.2< the weft density< 21.5) and (the warp yarn number > 9.5) and (the weaving type 2/1 Z) and (17< the weft yarn number <20), THEN the weight of the washing is low.

According to the rules, it has been observed that as the warp yarn number increases and the weft yarn number decreases, the weight of the washed denim fabric is low. If the weaving type is 2/1 twill, the weight of the washed denim fabric decreases. If the weave type is 3/1 twill the weight of the washed denim fabric increases. When the weft yarn material is mixed, the weight of the washed denim fabric increases. In addition, weft density is effective on the weight of the washed denim fabric.

In Table 8, the classification performance comparison results of the algorithms used for home type washing width of the denim fabric are given. Fabric width is an important quality parameter that affects purchasing factors, just like fabric weight, and is affected by many production parameters. The width per unit values of the denim fabrics have been measured after 3 household washing processes according to the standards. The rules between width and production parameters results have been presented in Table 9.

Table 9. *Results of classification algorithms for home type washing width*

	Accuracy	Precision	Recall	F Measure	Roc Area
Random tree	92.37%	0.92	0.92	0.92	0.97
J48	86.58%	0.86	0.86	0.86	0.94
Decision table	78.20%	0.78	0.78	0.77	0.92
Naive Bayes	61.73%	0.64	0.61	0.62	0.80
Simple logistic	62.34%	0.62	0.62	0.62	0.81
Hoeffding tree	60.67%	0.63	0.60	0.1	0.79
Suggested algorithm	96%	0.96	0.96	0.96	0.98

Table 9 demonstrates that the proposed algorithm yields the maximum result (96%) according to the accuracy, precision, sensitivity, F criterion, Roc Area performance metrics. It is followed by random tree, J48, decision table, naive Bayes, simple logistics algorithms, respectively. The Hoeffding tree algorithm yields a result of 60.67%. The extracted best rules for the weight of the denim fabric by the proposed algorithm are illustrated in Table 10.

Table 10. *Examples of extracted rules with production parameters for home type washing width of the denim fabric*

Rule 1	IF (the weft yarn material is 100% polyester) and (the weft yarn number>= 8) and (the weaving type is 3/1 Z) and (the warp yarn number <8.9), THEN the washing width of the denim fabric is high.
Rule 2	IF (the weft yarn material is 100% polyester) and (the weft yarn number>= 8) and (the weaving type is broken twill), THEN the washing width of the denim fabric is medium.
Rule 3	IF (the weft yarn material is mixture 95% tencel, 5% elastane or 96% cotton, 4% elastane) and (16.25 < the weft density <18.25) and (the warp yarn number>= 6.85), THEN the washing width of the denim fabric is low.

According to the rules, it is seen that as the weft yarn number increases, the washing width of the denim fabric increases. The weft yarn content is effective on the width of the washed denim fabric. In similar studies [41, 42], the effect of fiber properties and washing properties has been shown to be effective on fabric width.

5. RESULTS

It is a fact that the abundance of data is a problem not the scarcity of it, and the data storage and as the processing speed of computers increases, the currency of data mining, which is used to extract hidden patterns among data in bulk, is also increasing day by day. Among the data mining methods, classification that shows the information hidden in databases in the form of rule lists or decision trees that users can understand has become to draw due attention.

The current study investigated the link between denim fabric manufacturing factors and quality features. A novel classification rule extraction algorithm known as ABCLASS-miner is developed specifically for processing denim fabric production data to determine effective production parameters that cause different quality levels. The suggested rule extraction approach is based on the metaheuristic of artificial bee colony optimization. Experimentation and comparison studies on real fabric data have revealed that the suggested artificial bee-based algorithm is quite good in categorizing the quality parameters that are beneficial in denim fabric manufacture. The suggested method outperforms the other algorithms in terms of performance. As a result of the obtained rules, it is ensured that the enterprise has a foresight for the quality parameters.

ACKNOWLEGMENT

This study was supported by the Scientific and Technological Research Council of Turkey under Grant TUBITAK: TEYDEB-1505, 520006. Gözde Katırcıoğlu is supported as a YÖK 100/2000 scholar.

CONFLICTS OF INTEREST

No conflict of interest was declared by the authors.

REFERENCES

[1] Erdem, S., Özdağoğlu, G., “Analyzing of emergency data of a training and research hospital in aegean region using data mining”, Anadolu University Journal of Science and Technology, 9(2): 261-270, (2008).

[2] Frawley, W., Piatetsky-Shapiro, G., Maktheus, CW., “Knowledge discovery in databases: an overview”, AI Magazine, 13(3): 213–238, (1992).

[3] Han, J., Kamber, M., “Data mining: Concepts and techniques”, San Francisco, CA: Morgan Kaufmann Publishers, (2001).

[4] Yildirim, P., Birant, D., Alpyildiz, T., “Data mining and machine learning in textile industry”, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(1): e1228-e1248, (2017).

[5] Ozaydin, B., Hardin, J.M., Chhieng, D.C., “Data mining and clinical decision support systems”, Clinical Decision Support Systems, 45–68, (2016).

[6] Khedr, A. E., Idrees, A.M., Hegazy, A.E.-F., El-Shewy, S., “A proposed configurable approach for recommendation systems via data mining techniques”, Enterprise Information Systems, 12(2): 196-217, (2017).

[7] Stahl, F., Gabrys, B., Gaber, M.M., Berendsen, M., “An overview of interactive visual data mining techniques for knowledge discovery”, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 3(4): 239–256, (2013).

[8] Viloria, A., Acuña, G.C., Alcázar Franco, D.J., Hernández-Palma, H., Fuentes, J.P., Rambal, E.P., “Integration of data mining techniques to postgreSQL database manager system”, Procedia Computer Science, 155: 575–580, (2019).

[9] Chang, W., Hongjun, G., Rui, S., Huiling, S., “Technological potential analysis and vacant technology forecasting in the graphene field based on the patent data mining”, Resources Policy, 77 (1): 102636, (2022).

[10] Xiang, Q., Lv, Z.J., Yang, J.G., Yin, X.G., “Mining rule of quality control for spinning process with rough set theory”, Applied Mechanics and Materials, 80-81: 1021-1026, (2011).

[11] Deng, Z., Wang, L., Wang, X., “An integrated method of feature extraction and objective evaluation of fabric pilling”, The Journal of the Textile Institute, 102(1): 1-13, (2011).

[12] Kumar, K.V.N., Ragupathy, U.S., “An intelligent scheme for fault detection in textile web materials”, International Journal of Computer Applications, 46(10): 24-29, (2012).

[13] Xiang, Q., Lv, Z.J., Yang, J.G., “A novel data mining method on quality control within spinning process”, Applied Mechanics and Materials, 224: 87–92, (2012).

[14] Mozafary, V., Payvandy, P., “Application of data mining technique in predicting worsted spun yarn quality”, The Journal of The Textile Institute, 105(1): 100-108, (2013).

[15] Pang, KV., Chan, HL., “Data mining-based algorithm for storage location assignment in a randomised warehouse”, International Journal of Production Research, 55(14): 4035-4052, (2017).

[16] Mohanty, AK., Bag, A., “Detection and classification of fabric defects in textile using image mining and association rule miner”, International Journal of Electrical, Electronics and Computers (EEC Journal), 2(3): 28-33, (2017).

[17] Lizarraga-Morales RA., Correa-Tome FE., Sanchez-Yanez RE., Cepeda-Negrete J., “On the use of binary features in a rule-based approach for defect detection on textiles”, IEEE Access, 7: 18042-18049, (2019).

[18] Revathy, G., Gomathi, T., Sathish, E., “A comparative study for fault detection and classification in textile web materials”, Journal of Xi'an University of Architecture & Technology, 12(12), (2020).

[19] Das S., Ghosh A., “Rough set-based decision tool for classification of cotton yarn neps”, Journal of The Institution of Engineers (India), 102: 1–10, (2020).

[20] Bhuvaneshwarri, I., Tamilarasi, A., “Artificial intelligence and evolutionary computations in engineering systems”, Advances in Intelligent Systems and Computing, 1056: 69-77, (2020).

[21] Bhuvaneshwarri, I., Tamilarasi, A., “Predicting the fabric width of single jersey cotton knitted fabric using appropriate software”, Industria Textila, 70(6): 538-546, (2019).

[22] Deng, X., Zeng, X., Vroman, P., Koehl, L., “An intelligent multi-criteria optimization method for quick and market-oriented textile material design”, Journal of Global Optimization, 51(2): 227-244, (2010).

[23] Fu, Y., Chen, S., Chang, L., “Ant colony algorithm’s application of textile monitoring image recognition”, Advanced Materials Research, 328: 1701-1704, (2011).

[24] Zhao, X., Li, D., Yang, B., Ma, C., Zhu, Y., Chen, H., “Feature selection based on improved ant colony optimization for online detection of foreign fiber in cotton”, Applied Soft Computing, 24: 585-596, (2014).

[25] Xue, Z., Zeng, X., Koehl, L., Shen, L., “Development of an intelligent model to predict tactile properties from visual features of textile products”, 2015 10th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), (2015).

[26] Zhao, X., Liu X., Li, D., Chen, H., Liu, S., Yang, X., Zhan, S., Zhao, W., “Comparative study on metaheuristic-based feature selection for cotton foreign fibers recognition”, International Conference on Computer and Computing Technologies in Agriculture, 8-18, (2015).

[27] Amor, N., Noman, MT., Petru, M., Mahmood, A., Ismail, A., “Neural network-crow search model for the prediction of functional properties of nano TiO2 coated cotton composites”, Scientific Reports, 11(1): 13649, (2021).

[28] Periyasamy, A.P., Ramamoorthy, S.K., Lavate, S.S., “Eco-friendly denim processing”, Handbook of Ecomaterials, 1559–1579, (2019).

[29] Karthik, T., Murugan, R., “Carbon footprint in denim manufacturing”, Sustainability in Denim, 125–159, (2017).

[30] Üstündağ, S., “Development of elastic hybrid yarns that can be used in denim fabric production”, Master Thesis, Erciyes University Institute of Science and Technology, Kayseri, (2014).

[31] Dirican, A., “Evaluation and comparison of diagnostic test performance”, Cerrahpaşa Medical Journal, 32: 25-30, (2001).

[32] Niuniu X., Yuxun, L., “Review of decision trees”, The Third IEEE International Conference on Computer Science and Information Technology, 5: 105-109, (2010).

[33] Fan, W., Wang, H., Yu, P.S., Ma, S., “Is random model better? On its accuracy and efficiency”, The Third IEEE International Conference on Data Mining, 51-58, (2003).

[34] Duggal, P., Shukla, S., “Prediction of thyroid disorders using advanced machine learning techniques”, 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 670-675, (2020).

[35] Domingos, P., Hulten, G., “Mining high-speed data streams”, Proceedings of the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Boston, 71-80, (2000).

[36] Domingos, P., Pazzani, M., “Beyond independence: Conditions for optimality of a simple Bayesian classifier”, Machine Learning, 29: 103-130, (1997).

[37] John GH, Langley, P., “Estimating continuous distributions in Bayesian classifiers”, Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence (UAI'95): 338-345, (1995).

[38] Landwehr, N., Hall, M., Frank, E., “Logistic model trees”, Machine Learning, 59(1): 161-205, (2005).

[39] Ron K., “The power of Decision Tables”, European Conference on Machine Learning (ECML), (1995).

[40] Ghosh, A., Guha, T., Bhar, R.B., “Identification of handloom and powerloom fabrics using proximal support vector machines”, Indian Journal of Fibre and Textile Research, 40(1): 87-93, (2015).

[41] Ertaş, O. G., Zervent Ünal, B., Çelik, N., “Analyzing the effect of the elastane-containing dual-core weft yarn density on the denim fabric performance properties”, The Journal of The Textile Institute, 107(1): 116-126, (2016).

[42] Gürkan Ünal, P., Konal, D., “Investigation of the effect of weft yarn parameters on the elasticity and recovery properties of stretch denim fabrics”, Journal of Natural Fibers, 19(13): 7186-7198, (2021).