



Application of a newly developed naive Bayes algorithm in fire alarm

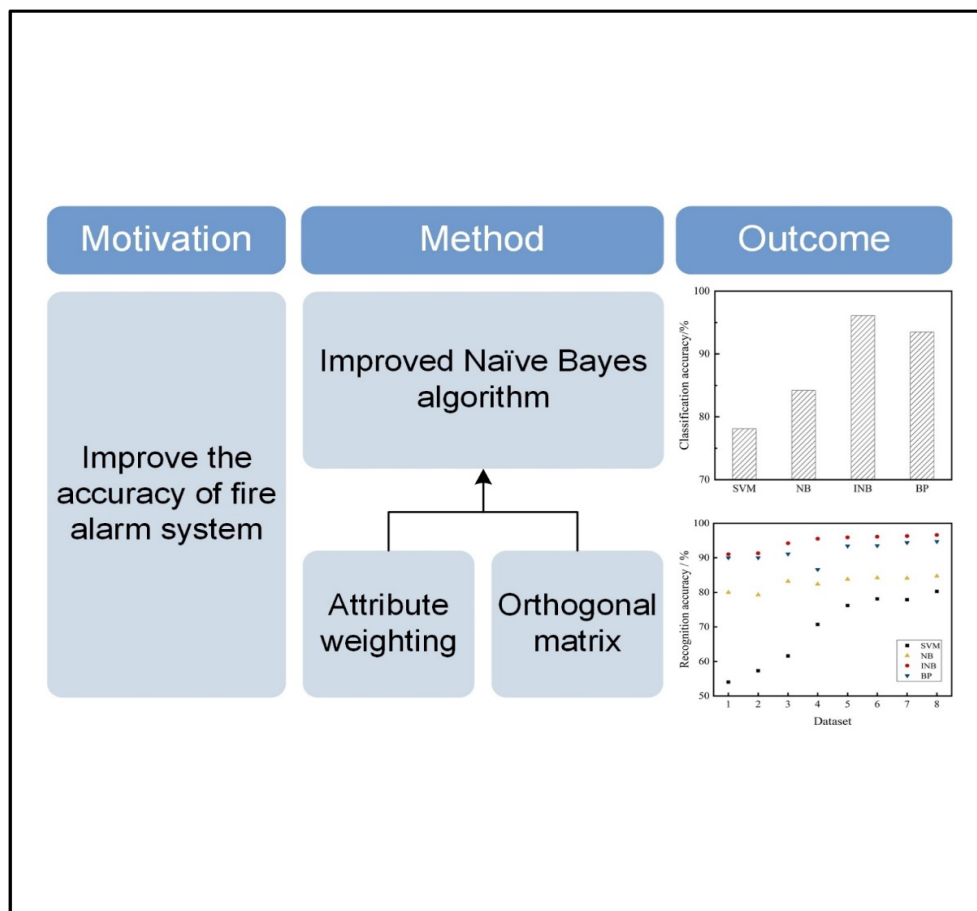
Xiangyong He, Yong Jiang , and Yong Hu

State Key Laboratory of Fire Science, University of Science and Technology of China, Hefei 230027, China

Correspondence: Yong Jiang, E-mail: yjjiang@ustc.edu.cn

© 2022 The Author(s). This is an open access article under the CC BY-NC-ND 4.0 license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Graphical abstract



With reasonable improvements, the naive Bayes algorithm can be the core processing algorithm of a fire alarm system.

Public summary

- A fire alarm algorithm based on naive Bayes was proposed.
- Attribute weighting and orthogonal matrix methods were introduced to improve naive Bayes.
- The improved naive Bayes algorithm has better performance and does not rely on a large amount of training data.

Application of a newly developed naive Bayes algorithm in fire alarm

Xiangyong He, Yong Jiang , and Yong Hu

State Key Laboratory of Fire Science, University of Science and Technology of China, Hefei 230027, China

✉Correspondence: Yong Jiang, E-mail: yjjiang@ustc.edu.cn

© 2022 The Author(s). This is an open access article under the CC BY-NC-ND 4.0 license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



Cite This: *JUSTC*, 2022, 52(6): 5 (8pp)



Read Online

Abstract: To address the problems of low recognition accuracy of traditional early fire warning systems in actual scenarios, a newly developed naive Bayes (NB) algorithm, namely, improved naive Bayes (INB), was proposed. An optimization method based on attribute weighting and an orthogonal matrix was used to improve the NB algorithm. Attribute weighting considers the influence of different values of each attribute on classification performance under every decision category; the orthogonal matrix weakens the linear relationship between the attributes reducing their correlations, which is more closely related to the conditional independence assumption. Data from the technology report of the National Institute of Standards and Technology (NIST) regarding fire research were used for the simulation, and eight datasets of different sizes were constructed for INB training and testing after filtering and normalization. A ten-fold cross-validation suggests that INB has been effectively trained and demonstrates the stable ability in fire alarms when the dataset contains 190 sets of samples; namely, the INB can be fully trained by using small datasets. A support vector machine (SVM), a back propagation (BP) neural network, and NB were selected for comparison. The results showed that the recognition accuracy, average precision, average recall, and average F_1 measure of INB were 96.1%, 97.3%, 97.2%, and 97.3%, respectively, which is the highest among the four different algorithms. Additionally, INB has a better performance compared to NB, SVM, and BP neural networks when the training time is short. In conclusion, INB can be used as a core algorithm for fire alarm systems with excellent and stable fire alarm capabilities.

Keywords: improved naive Bayes; conditional independence assumption; small sample data; recognition accuracy; fire warning

CLC number: X932

Document code: A

1 Introduction

Fire accidents remain to occur frequently in recent years and endanger human lives and properties owing to the high temperatures, smoke, and toxic gases produced by combustion^[1]. Fire alarm systems are an effective means of early alarm and fire control. With the timely activation of fire alarm systems, casualties and economic losses can be reduced, in addition to preventing fire accidents from spreading exponentially^[2]. Therefore, an accurate warning of fire accidents is critical for the safety of living beings, production, and emergency relief work.

Fire alarm systems are critical for firefighting and reducing damage caused by fires. They are installed to detect a critical fire situation and sound an alarm when a fire is detected^[2,3]. These systems consist of the following two features: fire detection, and processing of the detected characteristic data. The former refers to the performance of the sensors used to detect the characteristic data, which is the premise of a fire alarm. The latter refers to the relevant processing algorithm, which is at the core of the system. The input characteristics for data processing directly affect the final decision of the entire system and are critical for determining whether a fire

alarm system is adequate.

The early fire alarm system consisted of mainly a single sensor. This detection method, which considered only the temperature or smoke threshold as the basis of decision, is simple in structure and convenient to operate; however, it is vulnerable to interference and cannot accurately detect complex fire scenarios. Photoelectric smoke detectors^[4] emerged during the development of photoelectric technology in the late 1970s, and their anti-interference ability and stability were improved to a certain extent. In the early 1990s, inspiratory fire detection systems using laser technology significantly improved the sensitivity of fire detection. With the recent rapid development of computer and artificial intelligence technologies, fire alarm systems have also entered the intelligent stage. A fire alarm algorithm has been widely considered by researchers as the core of fire alarm systems and certain research achievements were obtained.

An artificial neural network (ANN) is used as a common method for fire prediction. Wu et al.^[5] utilized temperature, smoke concentration, and carbon monoxide as the input data to predict fires; a back propagation (BP) neural network was chosen to combine these three characteristics. Barera et al.^[6] combined neural networks with fuzzy reasoning and pro-

posed an intelligent fire alarm system that identifies fires in addition to informing individuals of these fires. Saeed et al.^[7] provided a new convolutional-neural-network-based model for early fire detection. The accuracy of the model is greater than 99%, which can be increased following more training. However, this method typically has a complex network structure and several parameters, which makes it difficult to achieve an embedded system. Additionally, implementing ANN usually requires several calculations, resulting in a high hardware cost^[8]. Therefore, this method is not suitable for fire alarm systems.

Some machine-learning algorithms with relatively simple structures and few parameters have been studied and compared. Bake et al.^[1] proposed a framework to identify fires based on a support vector machine (SVM) with a dynamic time-warping kernel function. This new framework achieves improvements in terms of fire detection time and false alarm rate; however, it does not consider the correlation between the time-series data. Kuo et al.^[9] studied a fire alarm device based on a gray-fuzzy algorithm to improve the traditional ship fire alarm system. The final test results indicate that the gray-fuzzy algorithm, which combines fuzzy rules and the gray theory, is feasible for real-time fire detection. Wei et al.^[10] proposed a fire alarm system that adopts the naive Bayes (NB) algorithm as a core algorithm for sensor information fusion. The simulation results demonstrated that the reliability of the system was improved. Additionally, comparative studies and discussions have been conducted on commonly used machine learning algorithms for data processing in fire alarm systems. Sulistian et al.^[11] designed a fire early warning system and implemented the following four machine learning algorithms for comparison: NB, SVM, decision tree, and K-nearest neighbor. A comparison of the results demonstrates that NB achieves the best performance in fire prediction. Because sensor nodes require a computationally cheap yet efficient algorithm to conduct fire detection in a near real-time manner, Bahrepour et al.^[12] proposed the use of NB and the feed forward neural network (FFNN) and introduced these two algorithms into a wireless sensor network to alarm for early residential fires. Comparative experimental results indicate that the NB classifier can achieve a better accuracy and has a lower communication overhead. Based on previous studies, it is reasonable to conclude that the NB has advantages including the simple programming to a sensor node, low computational cost, and high recognition accuracy, which makes it a promising algorithm for data processing in fire alarm systems.

The NB is a probability-theory-based method that has been widely used in applications such as spam email recognition^[13] and text emotion analysis^[14]. The calculation in NB is based on the assumption that all attributes are fully independent of one another given the class, which is called the conditional independence assumption. Apparently, the conditional independence assumption in NB is often violated in reality, which results in suboptimal probability estimates^[15] and harms its performance in applications with complex attribute dependencies^[16]. Certain reasonable improvements must be made to alleviate the conditional independence assumption. For NB with discrete attributes, Li et al.^[17] innovatively proposed the orthogonal transformation of discrete attributes with an ortho-

gonal matrix, which enhanced the independence between attributes to make it more in line with the conditional independence assumption. The comparative research results demonstrate that the classification performance of the improved algorithm after the orthogonal transformation was significantly improved. Attribute weighting is another commonly used approach for improving the NB. Shu et al.^[8] developed a NB algorithm using a double-weighting method for fire alarms. Experiments demonstrate that the algorithm improves the identification accuracy; however, the experimental data are not sufficiently comprehensive, and only four combustible materials are involved. Jiang et al.^[16] proposed a class-specific attribute weighted method and validated its effectiveness.

In this study, we developed a high-performance algorithm for fire alarm systems based on NB, named improved naive Bayes (INB), for which the attribute weighting method and orthogonal matrix method are introduced to weaken the conditional independence assumption in the traditional NB, which improves the recognition accuracy of the algorithm, reduces the false alarm rate, and increases the identification accuracy of a fire situation in the case of small data samples. By comparing INB with certain typical algorithms (SVM, NB, BP neural network), the following sections aim to illustrate its effectiveness and characteristics.

2 Naive Bayes and its optimization

2.1 Naive Bayes

For a test data sample \mathbf{x} , which is represented by an attribute value vector (a_1, a_2, \dots, a_n) in dataset \mathbf{D} , its class membership probabilities and class label are estimated by Eqs. (1) and (2), respectively, as follows:

$$\widehat{P}(c|\mathbf{x}) = \frac{\widehat{P}(c) \prod_{i=1}^n \widehat{P}(a_i|c)}{\sum_{c' \in C} \widehat{P}(c') \prod_{i=1}^n \widehat{P}(a_i|c')} \quad (1)$$

$$h_{nb}(\mathbf{x}) = \operatorname{argmax}_{c \in C} \widehat{P}(c) \prod_{i=1}^n \widehat{P}(a_i|c) \quad (2)$$

where n is the number of attributes, a_i is the i th attribute value of \mathbf{x} , and C is the collection of all class labels c .

The calculation of probabilities in NB is based on the conditional independence assumption; that is, all attributes are fully independent of one another in the given class. Although the conditional independence assumption reduces the computational cost, it is rarely true in reality, which may harm the performance of NB in circumstances with complex attribute dependencies^[16]. Attribute weighting and an orthogonal matrix are utilized to alleviate the conditional independence assumption in NB, fully retaining the advantage of NB while reducing the adverse effects of the assumption.

2.2 Improved naive Bayes

The linear relationship between the attributes is eliminated by an orthogonal matrix^[17] to reduce their correlations and improve the performance of the algorithm.

Let $\mathbf{D}_c = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ denote the set of all the samples belonging to class c in dataset \mathbf{D} , the sample $\mathbf{x}_k = (x_{k1}, x_{k2}, \dots, x_{kn})$ in \mathbf{D}_c is an n -dimensional vector. The covariance matrix was calculated as shown in Eq. (3):

$$M_c = \frac{1}{m} \sum_{k=1}^m (x_k - \mu)(x_k - \mu)^T \quad (3)$$

$$\mu = (\mu_1, \mu_2, \dots, \mu_n)^T \quad (4)$$

$$\mu_l = \frac{x_{1l} + x_{2l} + \dots + x_{ml}}{m}, l = 1, 2, \dots, n \quad (5)$$

where μ_l is the mean of the l th attribute value in each class, m is the number of samples in D_c , and M_c is the covariance matrix.

M_c is an $n \times n$ matrix, let the eigenvalues of M_c be $\lambda_1, \lambda_2, \dots, \lambda_n$, and the eigenvectors be $\beta_1, \beta_2, \dots, \beta_n$. Each eigenvector is unitized by Eq. (6) to obtain a standard orthogonal basis and to construct the orthogonal matrix P_c .

$$P_c = (\eta_1 \ \eta_2 \ \dots \ \eta_n) \quad (6)$$

in which, $\eta_i = \frac{1}{|\beta_i|} \beta_i, i = 1, 2, \dots, n$, where $\eta_i, i = 1, 2, \dots, n$, is the unitized eigenvector.

When the eigenvalue of M_c has repeated roots, the standard orthogonal basis can be obtained by Schmidt's orthogonal ligation and unitization to construct the orthogonal matrix P_c . After the covariance matrix M_c is harmoniously diagonalized by the orthogonal matrix P_c , all elements except the diagonal are zeros. Namely, the linear relationship between the attributes is removed, which is closer to the assumption of the conditional independence of NB.

The m samples in dataset D_c are transformed by the orthogonal matrix P_c , and the mean and variance of each attribute are then calculated. Following the aforementioned transformation, the mean is zero and the variance changes from $\sigma_{c,i}$ to $\sigma'_{c,i}$.

On this basis, the attribute weight $\delta_{c,i}$ is introduced to further optimize NB, and INB is expressed as follows:

$$h_{\text{inb}}(x) = \text{argmax}_{c \in C} \widehat{P}(c) \prod_{i=1}^n \delta_{c,i} \widehat{P}(y_i|c) \quad (7)$$

where $y = P_c^T(x - \mu_{c,i}), i = 1, 2, \dots, n$ is the new sample obtained by the orthogonal matrix transformation of sample x .

$$y = (y_1 \ y_2 \ \dots \ y_n),$$

$$\widehat{P}(y_i|c) = \frac{1}{\sqrt{2\pi\sigma'_{c,i}}} \exp\left(-\frac{y_i^2}{2\sigma_{c,i}'^2}\right),$$

$$\delta_{c,i} = \left[1 + \exp\left(-0.3 \times \frac{\prod_{c' \in C} f(\mu_{c,i}, \mu_{c',i})}{\sigma_{c,i}}\right)\right]^{-1},$$

$$f(\mu_{c,i}, \mu_{c',i}) = \begin{cases} |\mu_{c,i} - \mu_{c',i}|, & c \neq c'; \\ 1, & c = c'. \end{cases}$$

Here, c and c' are the class labels, $\mu_{c,i}$ and $\mu_{c',i}$ are the means of the i th attribute value in classes c and c' respectively, and $\sigma_{c,i}$ is the variance of the i th attribute value in class c .

The weight coefficient $\delta_{c,i}$ in INB is based on the weighting of classes and attributes. By analyzing the continuous data distribution of attributes in each class, the weight is constructed by using the variance of attributes and the mean of the product between the attributes. Among these, the attribute variance $\sigma_{c,i}$ can demonstrate the concentration of data, reflecting the classification superiority of the attributes under

this class while reducing the interference of the noise. A smaller attribute variance indicates a more centralized attribute data and is more conducive to the classification of samples. The product of the absolute value of the mean difference between different classes in the same attribute, namely $\prod_{c' \in C} f(\mu_{c,i}, \mu_{c',i})$, reflects the classification performance of the attribute; a better classification ability of an attribute is obtained when this value is larger.

3 Preparation for training data

Data from the home smoke alarm test report^[18] of the National Institute of Standards and Technology (NIST) were used for model training and testing. Full-scale tests in homes were performed during the experiments. Temperature, smoke, heat, carbon monoxide, and other data were recorded in detail during the tests, which provided meaningful and reliable data for the different stages of fire. Fourteen tests, which included tests 1 to 2, tests 4 to 8, test 10, test 11, test 15, test 33, test 35, test 38 to 39, were selected from the experiment, and data regarding temperature, carbon monoxide concentration, and smoke concentration were collected to construct eight datasets of different sizes for model training and testing. The partial data for the datasets are shown in Table 1. The following three different classes are identified in the table: no fire (NF), smouldering fire (SF), and open fire (OF). The impact of the dataset size on the model performance is discussed in Section 3.3.

3.1 Fire characteristic parameters

The combustion phenomenon and products may vary based on the environment or different stages of fire; however, similar characteristics remain to exist. A fire alarm system relies on the detection and processing of fire characteristic parameters^[19]. Because fire alarm systems focus on the identification of the initial stage of a fire, to achieve this function, it is

Table 1. Partial data of data sets.

Temperature	Smoke concentration	Carbon monoxide concentration	Class
0.00298	0.00004	0.13683	NF
0.00313	0.00004	0.15792	NF
0.85851	0.88500	0.25782	OF
0.06108	0.02562	0.34108	SF
0.00372	0.00000	0.00000	NF
0.85347	0.35861	0.00223	OF
0.00482	0.00044	0.10502	NF
0.17581	0.10198	0.25392	SF
0.01777	0.00953	0.11059	SF
0.18334	0.99870	0.01700	OF
0.95055	0.44156	0.00256	OF
0.00290	0.00025	0.00000	NF
0.01777	0.00953	0.11059	SF
0.59156	0.16889	0.00153	OF
0.00889	0.01590	0.01062	SF

necessary to reasonably select the number and types of fire characteristic parameters, which can accurately describe the spatial environment. The following three fire characteristic parameters are used in this study:

Temperature: The temperature constantly changes over time during the combustion process. In the early stage of a fire, the temperature is relatively low and slightly changes. When the fire stage changes from a smoldering to an open fire state, the temperature increases, which reflects the combustion state to a certain extent. Temperature is also one of the most easily measured indicators, thus can be utilized as an important characteristic parameter of fire.

Carbon monoxide concentration: Generally, carbon monoxide is nearly absent in air; however, its content rapidly increases when a fire occurs^[20]. Insufficient combustion produces carbon monoxide, especially during smoldering. As a unique characteristic of early fires, carbon monoxide has a low density, which easily floats to the top of an area and is easily detected by sensors. The concentration of carbon monoxide can be used as a reference for the fire stages.

Smoke concentration: Smoke mainly refers to the solid particles produced during the combustion process. This is one of the most apparent phenomena in a fire. Although it is not specific, it is one of the features of an early stage fire and is significant for characterizing the fire status^[21].

In summary, the INB considers temperature, carbon monoxide concentration, and smoke concentration as the three attributes, and outputs the three aforementioned fire state decision categories of NF, SF, and OF.

3.2 Data preprocessing

Because the fire characteristic parameters in the dataset are directly collected by the sensors, under normal circumstances, the parameters may produce random and irregular fluctuations owing to interference. However, when a fire occurs, the parameters exhibit an apparent and continuous trend of characteristics. Therefore, the filtering and noise reduction of data is required. The Savitzky-Golay (S-G) filter^[22] is a filtering method based on polynomial least-squares fitting in the time domain, which is widely used for data flow smoothing and noise reduction. S-G filtering is a weighted averaging algorithm for moving windows; however, its weighting coefficients are not a simple constant window but are obtained by least-square fitting of a given high-order polynomial in a slid-

ing window. The main characteristic feature of this filter is that the shape and width of the signal remains unchanged while filtering the noise. Smooth filtering with the S-G method improves data smoothing and reduces noise interference. A comparison of the carbon monoxide concentration data before and after noise reduction is shown in Fig. 1. According to Fig. 1a, there are many noise components in the raw data owing to the environment or electromagnetic interference; however, the overall characteristic trend is apparent. Fig. 1b presents the carbon monoxide concentration data after noise reduction when the window length is 41. Compared to the raw data shown in Fig. 1a, the noise of the data flow can be effectively reduced with the S-G filter.

To avoid the impact of certain attributes with large values (e.g. temperature) on smaller values (e.g. carbon monoxide concentration), all attribute values are normalized by using Eq. (8) and are mapped to the (0, 1) interval:

$$X_{\text{new}} = \frac{X_{\text{old}} - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \quad (8)$$

where X_{new} is the normalized data, X_{old} is the data prior to normalization, and X_{min} and X_{max} are the minimum and maximum values in the data, respectively.

As indicated in Eq. (8), if the noise reduction processing is not performed, the maximum and minimum values in the data are very likely to be noise values, which will affect the normalization results. Therefore, it is necessary to filter the noise-reduction process of the data before normalization.

3.3 Impact on data size

As a machine learning algorithm, the performance of INB has a strong relationship with the training set. The model cannot be fully trained when the dataset is small. However, in practical application scenarios, there may be problems, such as insufficient or missing data, resulting in a small dataset. Therefore, the effect of data size on the model performance needs to be considered. As previously indicated, NB is an algorithm that can handle classification tasks, performs sufficiently on small-scale datasets, and is not sensitive to missing values. Eight groups of datasets with different sizes were constructed to analyze the impact of dataset size on model performance.

Datasets 1 to 8 contain 100, 150, 190, 290, 580, 1150, 1530, and 2300 sets of samples, respectively. Similarly, the

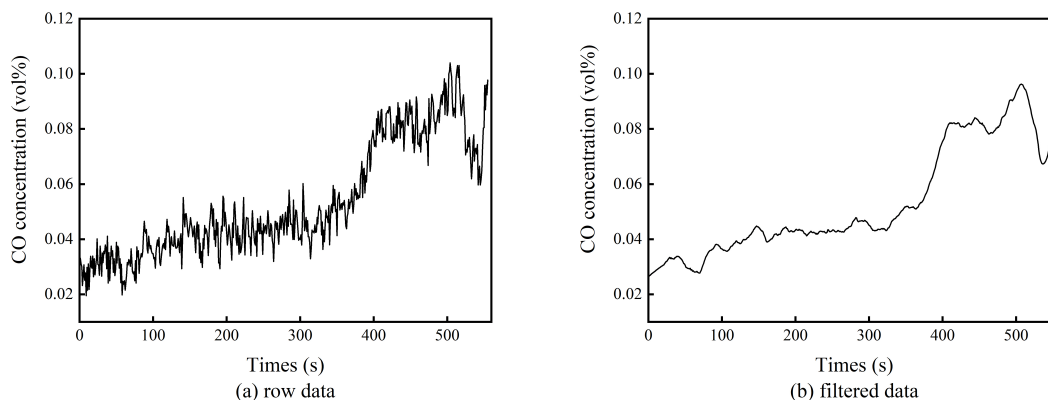


Fig. 1. Comparison of the data before and after filtering.

SVM algorithm, which also has advantages in dealing with small-scale datasets, is presented for comparison. The BP neural network is a commonly used method in the field of fire alarm algorithms and presents a good ability; thus, it is also presented. The specific parameter settings of the SVM and BP neural networks are described in Section 5. A ten-fold cross-validation^[16] was performed in the training and testing procedures, which divides dataset *D* into ten mutually independent subsets of a similar size, each of them trying to maintain consistency in data distribution. We then used the union of nine subsets as the training set and the remaining subset as the test set. In this manner, ten groups of training/test sets were obtained for ten training and testing times. The final return is the average of the ten test results. The classification accuracy of the four algorithms for different sizes of datasets is provided in Fig. 2, and its calculation formula is expressed by Eq. (9) in Section 4. The accuracy of each algorithm increases as the number of samples in the datasets increases and gradually tends to be stable; the classification results of the BP neural network is an exception in Dataset 4, which may be due to the design of the training data being inappropriate and resulting in a model that is not fully trained. The classification accuracies of the SVM, NB, INB, and BP neural network were stable at approximately 80%, 84%, 96%, and 94%, respectively. NB and INB achieved the classification accuracy in 190 samples (Dataset 3), whereas the BP neural network and SVM achieved the classification accuracy in 580 samples (Dataset 5) and 1150 samples (Dataset 6), respectively. The probability-theory-based NB and INB algorithms apparently have a greater advantage in dealing with small datasets with only approximately 190 samples, and the model can be sufficiently trained. The stable classification accuracy of INB is improved with attribute weighting and an orthogonal matrix by approximately 12% compared to that of NB. INB presents a good ability in fire classification accuracy; on the other hand, INB can be sufficiently trained and has a stable performance despite managing cases with small datasets. Thus, it does not depend on a large amount of training data and can be applied in cases of insufficient or missing data samples.

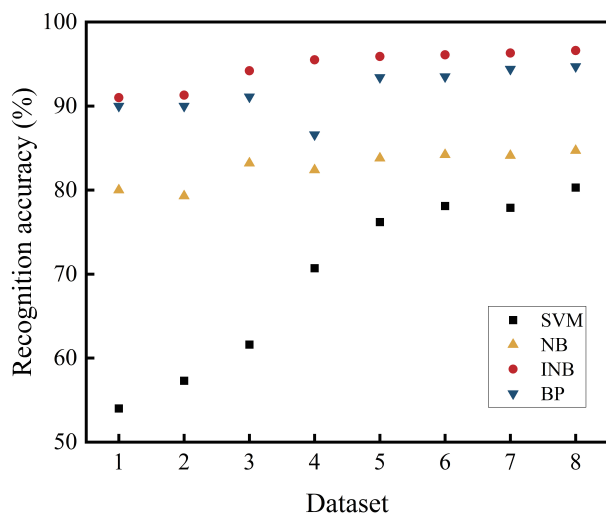


Fig. 2. Classification accuracy of four algorithms under different size of dataset.

4 Evaluation index

For binary classification problems, the classification results can be represented by the confusion matrix shown in Table 2, where TP, TN, FP, and FN indicate the corresponding sample numbers. The results of the multi-classification problems can also be expressed by a similar confusion matrix. Here, based on the confusion matrix, we used the following four indices to evaluate the performance of the algorithms: classification accuracy, precision, recall, and F_1 measure.

Classification accuracy is the proportion of the number of samples correctly identified by the model for all the samples. The calculation formula is given by the following:

$$ClaAcc = \frac{r}{R} \tag{9}$$

where ClaAcc is the classification accuracy, r is the number of samples correctly classified, and R is the total number of samples.

Precision is the proportion of the number of fire state samples correctly identified to the number of all such fire state samples identified by the model. The calculation formula is:

$$Pre = \frac{TP}{TP + FP} \tag{10}$$

where Pre is the precision of the model.

Recall is the proportion of the number of fire state samples correctly identified to the actual number of such fire state samples. The calculation formula is:

$$Rec = \frac{TP}{TP + FN} \tag{11}$$

where Rec is the recall of the model.

The F_1 measure considers the results of precision and recall, which is:

$$F_1 = \frac{2 \times Pre \times Rec}{Pre + Rec} \tag{12}$$

where F_1 is the F_1 measure of the model.

5 Comparative analyses and discussion

Based on the discussion in Section 3.3, INB can be concluded to have a good ability for managing small datasets. Furthermore, the advantages of INB over SVM, NB, and BP neural networks are discussed below, particularly in terms of classification accuracy and computational cost. The specific comparative analyses between these are as follows.

First, the four aforementioned algorithms were applied to the same scenarios and their results were compared. The penalty coefficient of the error item was 1 in the SVM, and the Gaussian kernel function was used. A penalty coefficient was

Table 2. The confusion matrix.

Actual results	Predicted results	
	Positive	Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)

used to control the penalty coefficient of the loss function. A larger penalty coefficient indicates a greater punishment for the wrong samples, thus leading to a higher accuracy in the training samples. However, the generalization ability of the model is reduced; that is, the classification accuracy of the test data is reduced. In contrast, if the penalty coefficient is reduced, some misclassification samples are allowed in the training samples, which has a strong generalization ability. The latter is generally used for training samples with noise, and incorrectly classified samples in the training sample set are regarded as noise. For the kernel function, when the data is inseparable in low dimensions, SVM maps the data to a high dimension through a Gaussian kernel function to make it separable in high dimensions, which solves the optimization problem in SVM. A BP neural network with two hidden layers was constructed for comparative analyses. Theoretically, a deeper number of hidden layers would lead to a stronger ability to fit the function, and subsequently better model results. However, a deeper number of hidden layers may cause overfitting, increase the difficulty of training, and make the model difficult to converge; therefore, a BP neural network with two hidden layers is constructed here. For neurons in hidden layers, not using enough neurons in the hidden layer will lead to underfitting. In contrast, when there are too many neurons in the hidden layer, the limited amount of information contained in the training set is insufficient to train all the neurons in the hidden layer, which leads to overfitting. Despite the training data containing sufficient information, too many neurons in the hidden layer will increase the training time; therefore, it is difficult to achieve the expected effect. Thus, the first and second hidden layers contained 100 and 50 neurons in our constructed BP neural network, respectively, and the maximum number of iterations was 500. The activation function adopts a rectified linear unit (ReLU) function. The role of the activation function is to increase the nonlinearity of the neural-network model. Otherwise, regardless of the number of layers in the neural network, the output is a linear combination of inputs, which is equivalent to the result of no hidden layers. In addition, using the ReLU function as an activation function solves the gradient vanishing problem while accelerating the convergence of the model. For NB, the probability density function of the normal distribution was used to calculate the probability. Dataset 6 was used here, which contained a total of 1150 samples, including 400 NF, 414 SF, and 336 OF. All samples were randomly disordered at first, and then divided into training and test sets according to the ten-fold cross-validation rules.

5.1 Classification accuracy

Classification accuracy can intuitively reflect the performance of the model. As shown in Fig. 3, after the improvement of attribute weighting and the orthogonal matrix, the classification accuracy of INB is the highest among the four algorithms, reaching 96.1%. Compared to NB, it improved by 11.9% mainly because the orthogonal matrix weakens the linear relationship between the attributes and reduces their correlations, which is more in accordance with the conditional independence assumption. On the other hand, the attribute weighting considers the influence of different values of each attribute on the classification performance under every decision category. The classification accuracy of SVM was the

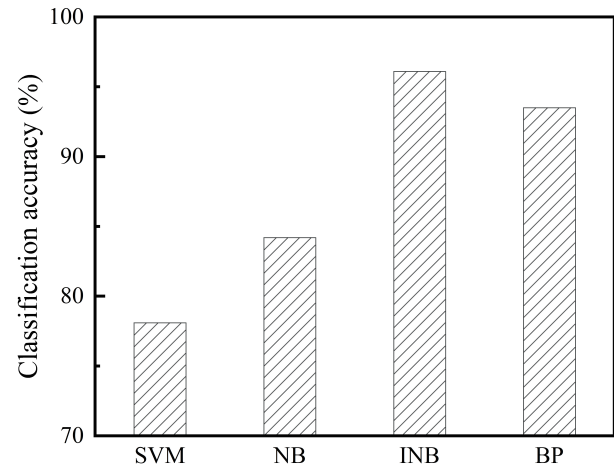


Fig. 3. Comparing classification accuracy of four algorithms.

lowest, only 78.1%. Although SVM has no local minimum problem compared with the BP neural network and can solve classification problems under small datasets, SVM has no general solution to nonlinear problems and is sensitive to missing data.

This may lead to a poor ability of the SVM in fire alarms. The classification accuracy of the BP neural network was second to that of INB, reaching 93.5%. However, owing to the existence of the feedback mechanism, the training time of the BP neural network was significantly longer than that of the other algorithms. A discussion of the computational cost is presented in Section 5.5.

Regarding the scenario of early fire alarms, the classification accuracy alone, as the evaluation index, does not fully demonstrate the ability of the model. For example, for 90 groups of no-fire conditions and 10 groups of fire conditions, the model can correctly classify all no-fire conditions, and all 10 groups of the fire conditions are incorrectly classified. The classification accuracy of this model was 90%. However, all the fire conditions are misclassified and do not play a role in fire alarms. Therefore, precision and recall were introduced to comprehensively evaluate the ability of the model.

5.2 Precision

A comparison of the precision results of the four algorithms is presented in Fig. 4. INB maintained the highest accuracy in all types of fire states, which increased by 2.2%, 16.1%, and 11.5% compared to that of NB for the three different fire states, respectively. The precision of NB for the SF state was low, only 79.1%. Following an improvement by utilizing the attribute weight and orthogonal matrix, the precision of the INB for the SF state increased to 95.2%; the precision of the SVM and BP neural network for the SF state were also low. The main advantage of INB is apparently the identification of the SF state. After the optimization, INB exhibited an excellent performance for all types of fire states, all of which reached more than a 95% precision. For the NF state, a SVM precision of 66.7% was obtained, which was well below the average, and only an 87.5% precision was obtained for the BP neural network. For the OF state, the precision of INB was lower than that of the BP neural network, with a gap of 3.3%. However, based on the discussion of the running times for the four algorithms in Section 5.5, INB has a notable advantage

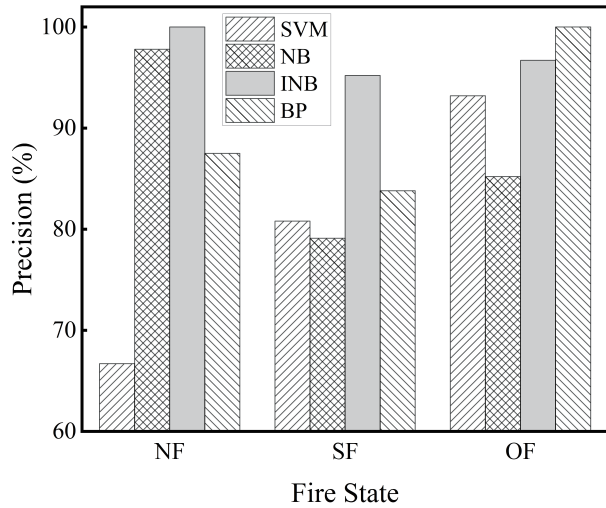


Fig. 4. Comparing precision of four algorithms.

over BP neural networks in terms of time complexity when the precision of the OF state does not lag significantly.

5.3 Recall

A recall comparison of the four algorithms is shown in Fig. 5. The INB recall reached approximately 95% in all three fire states, indicating an excellent performance. NB has a poor ability for the SF and OF states and has a low recall for both types of fire states. The misclassified samples of the SVM were mainly in the SF state. The recall value of the BP neural network for the three fire states is approximately 90%, which is relatively balanced but not as good as that of INB.

5.4 F₁ measure

As shown in Eqs. (10) and (11), precision and recall are a pair of contradictory measures. Generally, when the precision is high, the recall is often low, and vice versa. These two values can be simultaneously high only when the model presents an excellent ability and performs sufficiently well. According to Eq. (12), the F₁ measure is actually the harmonic mean of the precision and recall. Table 3 summarizes the results of the F₁ measure, which more intuitively presents a comparison of the four algorithms. The disadvantage of the poor recognition ability of NB on SF and OF is more apparent, as shown in Table 3. The F₁ measure of the SVM for OF was more than 90%, while the value was low for the other fire states. The performance of the BP neural network was better than that of NB and SVM, however, remained lower than that of INB.

A comparison of the results of the four aforementioned algorithms are summarized in Table 4. The average precision, average recall, and average F₁ measure are the averages of the precision, recall, and F₁ measure for each algorithm in the three fire states. As shown in the table, all of the INB evaluation indices reached more than 96%, which is significantly higher than that of the other algorithms.

5.5 Running time

The response speed of a fire alarm system is also an important indicator for fire alarms. Because some of the datasets were small, the separated training/testing time for those datasets was close to zero. Table 5 summarizes the average running times for the four algorithms on eight datasets with dif-

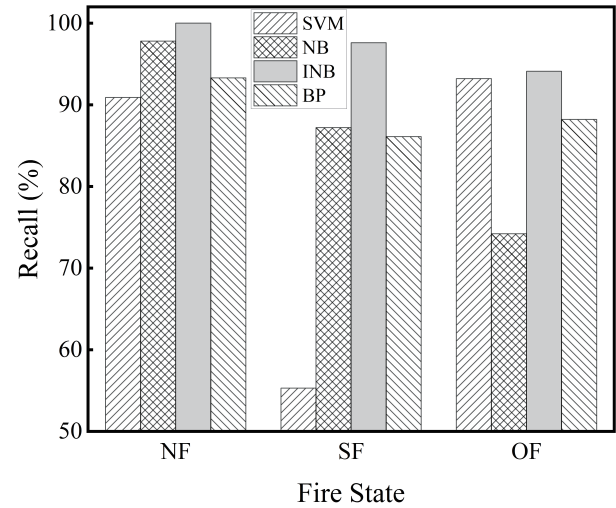


Fig. 5. Comparing recall of four algorithms.

Table 3. F₁ measure of four algorithms.

	SVM	NB	INB	BP
NF	76.9%	97.8%	100%	90.3%
SF	65.7%	82.9%	96.4%	84.9%
OF	93.2%	79.3%	95.4%	93.2%

Table 4. Summary test results of four algorithms.

Algorithms	Classification accuracy	Average precision	Average recall	Average F ₁ measure
SVM	78.1%	80.2%	79.8%	78.6%
NB	84.2%	87.4%	86.4%	86.7%
INB	96.1%	97.3%	97.2%	97.3%
BP	93.5%	90.4%	89.2%	89.5%

Table 5. Average running time on eight data sets of four algorithms.

	SVM	NB	INB	BP
Average training time (s)	0.01361	0.00335	0.06685	0.68236
Average testing time (s)	0.00124	0.00059	0.00439	0.00026

ferent sizes.

All the simulations were performed on the same computer (Intel Core Processor i5-6500 @ 3.20 GHz; 8.00 GB RAM; Windows 10 Pro., 64 Bit). Owing to the feedback learning processes present in the BP neural network, a longer training time is required. According to Figs. 4 and 5 and Table 5, although the BP neural network achieved a relatively high fire identification accuracy rate, an increase in the number of data samples was bound to affect the convergence speed. In addition, the complex network structure and significant number of parameters also make it difficult for the BP neural network to achieve an embedded system. SVM and NB have certain advantages in terms of training time; however, their classification accuracy is not high. Not only does INB have an excellent ability, but also it inherits the advantages of NB with respect to the training time, and can achieve a rapid response and accurate alarm in real scenarios.

6 Conclusions

As the core of the fire alarm system, the fire alarm algorithm not only has high requirements for fire recognition accuracy, but also requires it to be easily embedded. NB based on probability theory has a more stable and accurate classification performance, particularly when the data quantity is small^[15] and it is easy to achieve an embedded system. It has significant potential for the application of early fire alarms owing to its ability to manage uncertain evidence^[11]. However, for the traditional NB method, the conditional independence assumption limits its application for certain scenarios, and improvements to NB are therefore necessary to meet the aforementioned requirements of fire alarm algorithms. In this study, a newly developed algorithm, namely INB, based on NB, is proposed and applied to early fire alarms. Attribute weighting and an orthogonal matrix were used to improve the NB and weaken the conditional independence assumption. The INB was trained and tested using eight datasets of different sizes. To demonstrate the performance of INB in terms of classification accuracy and computational cost, it was compared with the SVM and BP neural networks. The major conclusions of this study are as follows:

(I) Comparative studies demonstrate that the classification accuracy, average precision, average recall, and average F_1 measure of the newly developed algorithm are significantly higher than those of the traditional NB, SVM, and BP neural networks, with specific values of 96.1%, 97.3%, 97.2%, and 97.3%, respectively.

(II) The calculation of the attribute weights and orthogonal matrices improved the recognition accuracy of INB in the fire states, while the testing time increased as a sacrifice, however remaining at a relatively low and acceptable level.

(III) The INB does not rely on a large amount of training data. It can also demonstrate a stable fire alarm ability when the quantity of data is small.

Acknowledgements

This work was supported by the Civil Aircraft Scientific Research Project of the Industry and Information Technology (BB2320000045, DD2320009001), and the Fundamental Research Funds for the Central Universities of China.

Conflict of interest

The authors declare that they have no conflict of interest.

Biographies

Xiangyong He is a postgraduate student under the supervision of Prof. Yong Jiang at the University of Science and Technology of China. His research mainly focuses on the application of machine learning in fire science.

Yong Jiang is a doctoral supervisor and professor at the University of Science and Technology of China, and is currently the director of Computer Simulation Research Office, State Key Laboratory of Fire Science. His research interests mainly include precision diagnostic experimental technology of fire and combustion, computer simulation and emulation of

fire and combustion, measurement and model of combustion reaction kinetics, thermal safety and artificial intelligence in energy utilization.

References

- [1] Baek J, Alhindi T J, Jeong M K, et al. Real-time fire detection algorithm based on support vector machine with dynamic time warping kernel function. *Fire Technology*, **2021**, 57 (6): 2929–2953.
- [2] Jafari M J, Pouyakian M, Khanteymooori A, et al. Reliability evaluation of fire alarm systems using dynamic Bayesian networks and fuzzy fault tree analysis. *Journal of Loss Prevention in the Process Industries*, **2020**, 67: 104229.
- [3] Shokouhi M, Nasiriani K, Cheraghi Z, et al. Preventive measures for fire-related injuries and their risk factors in residential buildings: A systematic review. *Journal of Injury & Violence Research*, **2019**, 11 (1): 1–14.
- [4] Chow W K, Wan E T K, Cheung K P. Possibility of using laser-fibre optics as a fire detection system. *Optics and Lasers in Engineering*, **1997**, 27 (2): 201–210.
- [5] Wu L S, Chen L, Hao X R. Multi-sensor data fusion algorithm for indoor fire early warning based on bp neural network. *Information*, **2021**, 12 (2): 59.
- [6] Sarwar B, Bajwa I S, Jamil J, et al. An intelligent fire warning application using IoT and an adaptive neuro-fuzzy inference system. *Sensors*, **2019**, 19 (14): 3150.
- [7] Saeed F, Paul A, Karthigaikumar P, et al. Convolutional neural network based early fire detection. *Multimedia Tools and Applications*, **2020**, 79 (13-14): 9083–9099.
- [8] Liang S, Zhang H G, You Y M, et al. Towards fire prediction accuracy enhancements by leveraging an improved naive bayes algorithm. *Symmetry*, **2021**, 13 (4): 530.
- [9] Kuo H C, Chang H K. A real-time shipboard fire-detection system based on grey-fuzzy algorithms. *Fire Safety Journal*, **2003**, 38 (4): 341–363.
- [10] Wei L M, Dong T H, Zhang Y X, et al. Research on fire alarm system based on bayesian algorithm. *Fire Science and Technology*, **2021**, 40 (8): 1199–1205.
- [11] Sulistian G, Abdurrohman M, Putrada A G, et al. Comparison of classification algorithms to improve smart fire alarm system performance. In: 2019 International Workshop on Big Data and Information Security (IWBIIS). IEEE, 2019: 119–124.
- [12] Bahrepour M, Meratnia N, Havinga P. Use of AI techniques for residential fire detection in wireless sensor networks. In: Proceedings of the Workshops of the 5th IFIP Conference on Artificial Intelligence Applications & Innovations (AIAI-2009), Thessaloniki, Greece, 2009: 311–321.
- [13] Ebadati O M E, Ahmadzadeh F. Classification spam email with elimination of unsuitable features with hybrid of GA-naive Bayes. *Journal of Information & Knowledge Management*, **2019**, 18 (1): 1950008.
- [14] Li Z, Li R, Jin G H. Sentiment analysis of danmaku videos based on Naive Bayes and sentiment dictionary. *IEEE Access*, **2020**, 8: 75073–75084.
- [15] Zaidi N A, Cerquides J, Carman M J, et al. Alleviating naive Bayes attribute independence assumption by attribute weighting. *Journal of Machine Learning Research*, **2013**, 24: 1947–1988.
- [16] Jiang L X, Zhang L G, Yu L J, et al. Class-specific attribute weighted naive Bayes. *Pattern Recognition*, **2019**, 88: 321–330.
- [17] Li F X, Wang J M, Liang J C, et al. Optimization of naive Bayesian classification algorithm for discrete attributes. *Journal of Chinese Computer Systems*, **2022**, 43 (5): 897–901.
- [18] Bukowski R W, Peacock R D, Averill J D, et al. Performance of home smoke alarms analysis of the response of several available technologies in residential fire settings. Gaithersburg, MD: National Institute of Standards and Technology, 2003.
- [19] Alessandri A, Bagnnerini P, Gaggero M, et al. Parameter estimation of fire propagation models using level set methods. *Applied Mathematical Modelling*, **2021**, 92: 731–747.
- [20] Zhang J J, Ye Z Y, Li K F. Multi-sensor information fusion detection system for fire robot through back propagation neural network. *PLoS ONE*, **2020**, 15 (7): e0236482.
- [21] Yang X, Zhang K, Chai Y, et al. A multi-sensor characteristic parameter fusion analysis based electrical fire detection model. In: Proceedings of 2018 Chinese Intelligent Systems Conference. Singapore: Springer, 2018: 397–410.
- [22] Savitzky A, Golay M J. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, **1964**, 36 (8): 1627–1639.