Journal of Theoretical and Applied Information Technology
30th April 2020. Vol.98. No 08

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

INDOOR FIRE DETECTION SYSTEM BASED ON DATA MINING TECHNIQUES

[1] ABDELWAHAB ALSAMMAK, [2] KHALED M. FOUAD, [3] SHAIMA D. ALMARIE

[1] Faculty of Engineering Shoubra, Benha University, Egypt

[2] Faculty of Computers and Informatics, Benha University, Egypt

[3] Faculty of Computers and Informatics, Benha University, Egypt

E-mail: [1]asammak@feng.bu.edu.eg, [2]kmfi@fci.bu.edu.eg, [3]galaty852012@gmail.com

ABSTRACT

Fire ignition is one of the events that can cause severe humanitarian, economic, and environmental damages. Fast fire detection can significantly help in controlling the fire and reducing the resulting losses. Efficient fire-fighting in schools can save many lives and resources. In this study, a data mining-based fire detection system is proposed. It can be used in detecting fire sources in different locations of Kuwaiti schools that are listed in the ministry of Education. The proposed system consists of four main steps: data acquisition, data preprocessing, feature analysis and selection, and classification. The data acquisition is performed with the help of the General Civil Defense Department of Kuwait. In the data preprocessing step, a set of operations is performed on the collected data, including discretization and categorization. In the feature selection step, two feature selection techniques are used, namely Information Gain (IG) and Principle Component Analysis (PCA). Finally, in the classification step, five classification models are used to perform the classification task, including Decision Tree (DT), Linear Regression (LR), Linear Discriminant (LD), Support Vector Machine (SVM), and Deep Belief Network (DBN). Intensive experiments are performed to evaluate the proposed system, and the obtained results are auspicious.

Keywords: *Fire Detection, Data Mining, Deep Learning, Classification.*

1. INTRODUCTION

Several abnormal events can happen in our lives, such as fire, accidents, floods, earthquakes, etc. These events can have catastrophic consequences on human beings/properties and the surrounding environment. Among such activities, fire is one of the commonly occurring events that can be effectively restrained during its early stages; and consequently, the resulting damages can be reduced [1]. Using fire has been one of the significant factors that helped to evolve the human civilization. However, out control fire can cause severe humanitarian, economic, and environmental damages that vary based on its type (causing factor), location, scale, etc. [2].

Fire is mainly the fast oxidation of a substance that yields gases and chemicals. Different sensors can read these chemical productions to give an insight into the type and place of the fire [3]. Active fire detection systems that can detect fire occurrence with minimum latency can play an essential role in preventing fire expansion and minimizing the resulting losses. Nowadays, traditional fire watchtowers are replaced automated fire detection systems that depend on different technologies including IR sensors [4], LIDAR (Light Detection and Ranging systems) [5], satellite platforms [6], to computer vision-based methods [7, 8] and WSN (wireless sensor network) systems [9]. In general, automatic fire detection systems can be categorized into three main classes [10]: satellite-based, infrared/smoke scanners, and local sensors (e.g., meteorological). Satellite-based systems have high acquisition costs, high latency, and wrong resolution in some cases, mainly indoor fires.

Additionally, the cost of the scanner-based system is relatively high due to the necessary equipment and maintenance costs [4]. Moreover, fire detection systems that depend on local sensors for capturing single fire characteristics such as temperature cannot achieve high detection accuracy. These systems cannot differentiate the temperature rise caused by fire occurrence from that caused by environmental changes; hence, these systems have a high false alarm/negative rate [5].

On the other hand, data mining (DM) techniques can be effectively employed in building intelligent fire detection systems. Data mining is an iterative process that aims at discovering valuable patterns and information from large datasets. Different data mining approaches have been successfully employed in many fields, such as business, healthcare, engineering, scientific research, etc. These approaches are originated from various disciplines, including statistics, artificial intelligence, machine learning, mathematics, and information theory. Several tasks can be performed using data mining techniques such as clustering, classification, association, and outlier detection. Hence, the data mining approach can be used to analyze the data collected using the different sensors and to predict the occurrence of fire [4, 11].

In this paper, a data mining-based fire detection system is introduced that can be used in detecting fire sources in different locations of Kuwaiti schools that are listed in the ministry of Education. The proposed system aims to automate the means of fire-fighting techniques and materials in order to systemize the process of both detecting the source of fire in any school by a far-distance system connected in the ministry administration building, and then determining the material of fire-fighting (water, rubber, bubble, etc.) in the method that suits both the fire point (class of students, a lap of computers, lap of chemical materials, etc.). This system is both vital and essential in the areas where precious lives on the stack. The proposed system consists of four main steps: data acquisition, data preprocessing, feature analysis and selection, and classification. The data acquisition is performed with the help of the General Civil Defense Department of Kuwait. In the data preprocessing step, a set of operations is performed on the collected data, including discretization and categorization. In the feature selection step, two feature selection techniques are used, namely Information Gain (IG) and Principle Component Analysis (PCA). Finally, in the classification step, five classification models are used to perform the classification task, including Decision Tree (DT), Linear Regression (LR), Linear Discriminant (LD), Support Vector Machine (SVM), and Deep Belief Network (DBN).

The remaining sections of the papers are arranged as follows: Section 2 reviews some of the research efforts that have been performed in fire-fighting based on the automatic fire detection system. Section 3 presents the proposed system with a detailed description of each step. Section 4 includes the implementation details and evaluation process for the proposed system. Finally, Section 5 contains the conclusion and future work.

2. RELATED WORK

Many efforts have been performed to develop smart fire detection and prediction systems. In this section, several works are reviewed and summarized.

Cortez and Morais [4] have proposed a forest fire prediction system based on meteorological data and several well-known classifiers. The proposed system aimed at predicting the burned areas. Different approaches have been adopted to select the essential attributes. The conducted experiments have shown that the best prediction results were obtained using a support vector machine (SVM) classifier and four attributes, namely, temperature, wind, relative humidity, and rain.

Angayarkkani and Radhakrishnan [12] have presented a forest fire detection system that depends on analyzing the spatial data. In the proposed system, the fire regions are detected by converting the RGB spatial data into XYZ color space and segmenting the results using anisotropic diffusion. Then, a trained neural network is used to identify the fire regions. The conducted experiments on a public dataset have revealed the efficacy of the proposed system.

Bahrepour et al. [3] have introduced a comprehensive study to analyze the performance of several data mining techniques for indoor and outdoor fire detection based on wireless sensor networks (WSNs). Besides, the conducted research aimed at identifying the correlation between the different attributes that can be used in fire detection. The obtained results have shown that their proposed fire detection system that depends on distributed neural network and naïve Bayes can achieve fire detection accuracy up to 81% and 92% for indoor and outdoor fires, respectively.

Li and Wu [2] have proposed a system for fire detection in digital images based on data mining approaches. The proposed system merges the fire color model and frequent pattern mining in the fire detection process. The obtained results have proved the superiority of the proposed method compared to other well-known fire detection method called the Çelik's method.

Li et al. [8] have suggested a fire detection system that depends on the orientation feature. The proposed system depends on computing the orientation feature for the different pixels through the consecutive frames of the surveillance video. Then, an

SVM classifier is used to the fire regions. Intensive experiments have been conducted, and the obtained results have shown that the proposed method is better than the state-of-art methods in terms of both fire detection accuracy and fire detection speed.

Maksimović and Vujović [11] have conducted a comparative study to evaluate the performance of several data mining approaches in fire detection based on WSN. The evaluated data mining approaches include the Fuzzy Unordered Rule Induction Algorithm (FURIA), OneR, decision tree, Naïve Bayes, and neural network classifiers. The experiments performed on three different datasets have shown that the FURIA classifier gives the best fire detection results.

Maksimović et al. [13] have conducted a study to improve the efficiency of indoor WSN-based fire detection systems. The objective of the study to determine the optimal places of fire heat detectors that can prolong the life of WSNs in case of fire ignition. Besides, it aims at identifying the minimum number of sensors that achieve full coverage for the monitored area.

Hamadeh et al. [14] have proposed a forest fire prediction system that depends on data mining approaches as well as meteorological data. The used data mining approaches include decision tree and beck-propagation neural network. The used meteorological attributes include temperature, relative humidity, wind speed, and daily precipitation. In the proposed work, the decision tree is employed to determine the importance of each attribute. The conducted experiments have shown that a 2-input neural network can achieve higher prediction accuracy compared to the 4-input neural network.

Garcia-Jimenez et al. [15] have introduced a generalization framework for the Mamdani inference algorithm to improve the performance of fuzzy logic-based systems. Their proposed employs overlap functions and overlap indices. The proposed method has been evaluated in the field of forest fire detection based on WSNs. The conducted experiments have shown that the proposed method has competitive performance.

Saoudi et al. [16] have presented a forest fire detection system that integrates the data mining techniques in the individual sensors of the clustered WSN. In the proposed system, each sensor can make a decision about the occurrence of fire ignition based on the integrated classifier. In the case of fire detection, the sensor node alerts the cluster head, which in turn passes this information to other cluster heads until it reaches the sink to start the fire-fighting procedure. The proposed system has been evaluated using the CupCarbon simulator, and the Naïve Bayes classifier is integrated into each sensor node. The conducted experiments have shown that the proposed system can achieve the right balance between energy consumption and accurate fire detection.

Schroeder et al. [7] have introduced an active fire detection system based on Landsat-8 Operational Land Imager (OLI). The proposed system employs infrared and near-infrared channels in the detection process. Besides, it adopts the multi-temporal analysis to enhance pixel classification accuracy. The obtained results have proved the efficacy of the proposed system.

Lin et al. [17] have proposed a forest fire prediction system using rechargeable WSNs based on fuzzy inference and big data analysis. In the proposed system, the weather information of the forest is collected. Then, the fuzzification is performed to the inputs with a weighted fuzzy reasoning algorithm. The conducted experiments have shown that the proposed system can play an essential role in forest fire prediction.

Mahmud at al. [18] has presented an early home fire detection system based on data mining techniques. The proposed system collects some information, including heat, smoke, and flame, and passes it to the data mining algorithm to decide about fire ignition. In the case of fire detection, the proposed system informs several parties simultaneously, including a fire department, ambulance service, and police station. The performance of many data mining techniques is assessed. The obtained results have shown that several data mining techniques can achieve high fire detection accuracy.

Muhammad et al. [1] have introduced a cost-effective fire detection system based on surveillance video. The proposed system employs a convolutional neural network (CNN) whose architecture inspired by the GoogleNet Architecture. The proposed architecture attempts to make a balance between efficiency and accuracy. The conducted experiments on several fire datasets have shown that the proposed system has a competitive performance compared to the state-of-art approaches.

Al_Janabi et al. [19] have conducted a study to evaluate the performance of various soft computing techniques in forest fire prediction. First, the frequent patterns are identified using Principle Component Analysis (PCA), and the fire regions are clustered using the Particle Swarm Optimization algorithm. In the second level of the proposed sys-

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

tem, five soft computing techniques are employed in parallel in the prediction process. The obtained results of the SVM predictor is better than other predictors.

Xie and Peng [20] have proposed a forest fire prediction system based on a tuned random forest approach. The objective of the proposed system was to predict the burned areas and to predict the ignition of large-scale fires. The obtained results have shown the superiority of the random forest approach compared to other regression models.

3. THE PROPOSED SYSTEM

In this section, a fire detection system based on data mining techniques is presented. The objective of the proposed system is to detect the fire ignition in different locations of Kuwaiti schools that are listed in the ministry of Education.

The system aims to automate the means of fire-fighting techniques and materials in order to systemize the process of both detecting the source of fire in any school by a far-distance system connected in the ministry administration building, and then determining the material of fire-fighting (water, rubber, bubble, etc.) in the method that suits both the fire point (class of students, a lap of computers, lap of chemical materials, etc.) As shown in Fig. 1, the proposed system consists of four main steps: Data acquisition, data preprocessing, feature selection, and classification. A detailed description is given for each step separately in the following subsections.

3.1. Data Acquisition

The data was collected from 2014 to Q2 2018 from the General Civil Defense Department, Ministry of Interior in the different sites of the Kuwaiti public schools. The dataset was collected from six governments in Kuwait. In the dataset, there are 22 attributes, including spatial and temporal attributes. Besides, two types of fire sensors, including interior and exterior sensors. Each sensor has two components that are affected directly by the weather conditions.

Four meteorological attributes are consisting of latitude code, longitude code, day, and month. Also, the collected attributes contain the identity code, location type, risk plan, exit path, damage severity, dies, infection, fire-fighting, the response variable, temperature, humanity, wind speed, and raining state. The collected dataset includes 485 records. The details of the used dataset are presented in Table 1.

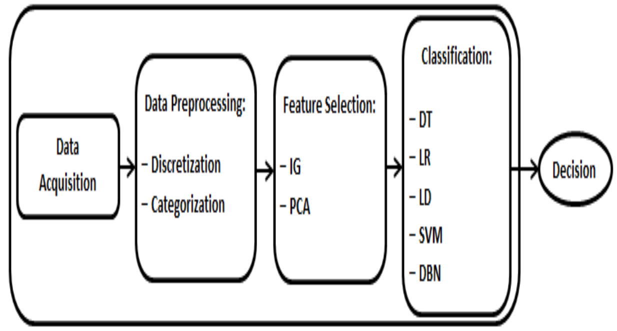

FIGURE 1. The block diagram of the proposed fire detection system.

Table 1: The Details of The Used Dataset

Attribute	Definition	Range
82.6	85.4	
83.1	82.6	
NO	Identity record for each case in the dataset	
MM	Represent the month in a year	01:12
DD	Represent the day in a month	01:31
Cause	Represents the physical issues that cause the fire in the school	A: professional negligence B: Smoking C: Electric D: Lab Experiment E: Children Riot F: Environmental factors G: Other no fire
Long.	Represents the X coordination of lookup map for region block's map	
Lat.	Represents the Y coordination of lookup map for region block's map	
Location Type	Represent the class of the fired block in the school	C: Class G: Garden L: laboratory O: Office No Fire Other
Temperature	Represents the normalized temperature in	

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Humidity	Represents the normalized humidity in	
Wind Speed	Represents the wind speed in the day of the fire event	
Raining	Represents the raining state in the day of the fire event	
Risk Plan	Identify if the school has a risk plan in case of fire or not	
Exit Path	Identify if the school has designed to have emergency exit path in case of fire or not	
S1A	Represent the first read of internal Fire Sensor	
S1B	Represent the second read of internal Fire Sensor	
Damage	Records the side effects of fire events	
Dies	Records the number of dies cases as the effects of fire events	
Infection	Records if there are infections cases the side effects of fire events	
S2A	Represent the first read of external Fire Sensor	
S2B	Represent the second read of external Fire Sensor	
Fire-fighting	Represents whether the fire has been fought by watch towers school staff or the defense staff.	
Fire	Represents the class label	

3.2. Data Preprocessing

In this step, two preprocessing operations are applied to the collected data. The first operation is the discretization in which all continuous-valued attributes such as the longitude, latitude, data collected by the sensors, temperature, humidity, and rain are discretized. The second operation is the categorization in which the symbolic class names are represented using numeric values. The set of attributes on which the categorization operation is applied include the cause, day, month, exit path, etc. For example, the class names for the cause attribute are represented using 1, 2, 3, etc. rather than A, B, C, etc.

3.3. Feature Selection

The objective of this step is to identify the most influential attributes to reduce the dimensionality of the input feature vector. Reducing the dimensionality of feature vectors has many advantages, such as accelerating the classification process and overcoming the over-fitting problem. Moreover, it can improve the classification accuracy. In this study, we have adopted two known feature selection/reduction algorithms: Principal Component Analysis (PCA) and Information Gain (IG). The ideas behind the used approaches are briefly described in the following subsections.

3.3.1. Principal component analysis

Principal component analysis (PCA) is a widely adopted statistical data analysis tool that can be employed to perform various tasks, including feature extraction, feature selection, classification, and data compression [21]. PCA generates pertinent features through linearly mapping correlated variables into a smaller group of uncorrelated variables, which are known as the main components or the principal components. The generated components are linear combinations of the original data that capture most of the variance in the data [22]. The PCA method can be summarized in the following steps [21]:

Step 1: The mean of the feature vectors is computed by (1).

$$m = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad (1)$$

Where denotes the number of samples.

Step 2: The covariance matrix is computed using (2).

Journal of Theoretical and Applied Information Technology
30th April 2020. Vol.98. No 08

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$C = \frac{1}{N}\sum_{i=1}^{N}(x_i - m)(x_i - m)^T \quad (2)$$

Step 3: The following decomposition is solved.

$$\lambda_i e_i = c e_i \quad (3)$$

Where λ_i is the eigenvalue associated with the eigenvector e_i.

Step 4: The transformation matrix is composed by selecting the first eigenvectors, which are ordered in a descending manner based on the corresponding eigenvalue, λ_i as shown in (4).

$$F = \{e_i\}_{i=1}^{g} \quad (4)$$

Step 5: The reduced feature vectors r_i can be obtained from the original vectors x_i using Eq. (5).

$$r_i = F^T x_i \quad (5)$$

3.3.2. Information gain

Information Gain (IG) is one of the simplest methods that can measure the dependence between the feature and the class label [23]. Hence, it is a widely adopted feature selection method due to its low computation overhead as well as being easy to interpret. The information gain of a specific feature X and the class label Y can be computed using (6) [24].

$$I(X,Y) = \eta(X) - \eta(X \mid Y) \quad (6)$$

Where entropy (η) is a measure of the uncertainty associated with a random variable $\eta(X)$

Moreover, $\eta(X \mid Y)$ is the entropy of the feature X and the entropy of the feature X after observing the class label Y, respectively. They can be computed as following [24].

$$\eta(X) = -\sum_i P(x_i)\log_2(P(x_i)) \quad (7)$$

$$\eta(X \mid Y) = -\sum_i p(x_i)\sum_i p(x_i \mid y_i)\log_2(p(x_i \mid y_i)) \quad (8)$$

The maximum information gain value is 1. The higher value means the higher dependence between the feature and the class label. The information gain is calculated independently for each attribute. Then, the attributes of higher information gains are used to compose the reduced feature vectors.

3.3.3. Classification

In this step, several classification models are used for fire detection. The used classifiers include Decision Tree (DT), Support Vector Machine (SVM), Linear Regression (LR), Linear Discriminant (LD), and Deep Belief Network (DBN). The ideas behind the different classifiers are shown below in the following subsections.

3.3.3.1. Decision trees

Decision Trees (DTs) are well-known and widely adopted tools for building classifiers in the field of data mining. These systems take a set of cases (Also, known as observations or patterns) as input where each case belongs to one of few classes and represented by a set of values corresponding to many attributes, and produce a classifier, which able to predict the class of new unseen cases, as output [25]. Decision tree learning is a standard method for inductive inference. It approximates discrete-valued functions that can handle noisy datasets and can learn disjunctive expressions. Besides, it can approximate discrete-valued target functions, where the produced functions are represented in the form of a decision tree. Moreover, the produced DT can be represented using a group of if-then rules for improving the readability by end-users [26]. A DT is a classifier which seeks for recursive portioning of the instance (pattern or data point) space. It consists of a root node, internal nodes, and leaves. Fig. 2 presents a decision tree that represents the possibility of responding to direct mailing by a customer, where circles represent the root and internal nodes and triangles represent the leaves.

A DT recursively divides the cases into branches based on several mathematical algorithms, such as information gain, Gini index, and Chi-squared test, to determine the attribute and corresponding threshold of that attribute that divides the input cases into two or more groups. This process is recursively performed at each leaf node until the tree construction is completed.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

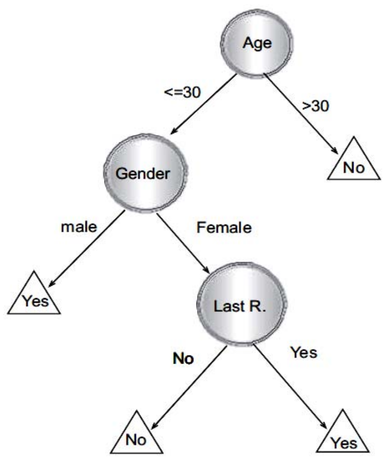

FIGURE 2. A decision tree represents the possibility of responding to direct mailing by a customer.

3.3.3.2. Support vector machine

Support Vector Machine (SVM) is a supervised learning technique that mainly presented to perform classification and regression tasks. It was proposed by Boser et al. and Vapnik [27]. It is utilized to determine the hyperplane, which isolates the classes in such a way that minimizes the training error and maximizes the margin to improve the generalization ability of the classifier [28].

Given a linearly separable data set, a linear SVM classifier can easily be employed for classifying it. The algorithm attempts to increase the margin among the classes. Support vectors are those points that exist on the margins, as demonstrated in Fig. 3 [28].

The primary duty of the generated hyperplane can be formulated using (9).

$$g(x) = w^T x + b \tag{9}$$

x describes data points, w denotes a coefficient vector, and b represents the offset from the origin point. When SVM is a linear algorithm, g(x) >=0 for the nearest point on the one that belongs to the class. On the other side, g(x) < 0 for the nearest point that belongs to another class. Margin ($2/\|w\|^2$) has to be as large as possible to improve the generalization capability, which in turn minimizes the cost function that is defined by (10) [28].

$$j(w) = \frac{1}{2}\|w\|^2 \tag{10}$$

Where $y_i(w^T x_i + b) >= 1$, i = 1, 2, …, n , and $y_i = \{+1, -1\}$ denotes class labels.

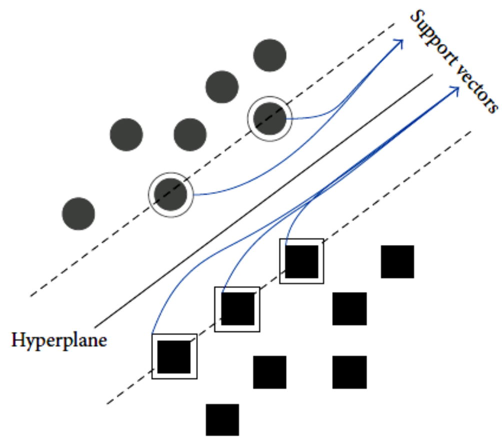

FIGURE 3. The separating hyperplane with support vectors [29].

The other usage of SVM is that it can solve nonlinear classification problems by employing a kernel function. The role of the used kernel function is to map the data points into another space of higher dimension to obtain a hyperplane that can separate the classes in hand. The new function of non-linear SVM can be formulated using (11) [28].

$$g(x) = w^T \Phi(x) + b \tag{11}$$

where $\Phi(x)$ denotes the input vectors' mapping onto the kernel space x.

3.3.3.3. Linear regression classification

Suppose that we have C classes, and each class has N samples in the n-dimensional feature space. Let, k =1, …, N be the n-dimensional feature vectors in the training set of the ith class. Then the predictor of the ith class is represented using (12) [30].

$$w_i = [w_1^i \vdots w_2^i \vdots \cdots \vdots w_N^i] \tag{12}$$

If a feature vector belongs to the ith class, then it can be represented by the linear combinations of these feature vectors with an error according to the linear regression classification [30]. Therefore,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$v = w_i \beta_i + e \quad (13)$$

where is the dimensional parameter vector. The sum of error squares can be computed using (14).

$$S = e^T e = (v - w_i \beta_i)^T (v - w_i \beta_i) \quad (14)$$

After the minimization concerning, the estimation of the vector parameters becomes:

$$\tilde{\beta}_i = (w_i^T w_i)^{-1} w_i^T v \quad (15)$$

Then the estimation of the vector can be computed using (16) [30].

$$\tilde{v}_i = w_i (w_i^T w_i)^{-1} w_i^T v \quad (16)$$

Here the projection matrix can be computed using (17).

$$M_i = w_i (w_i^T w_i)^{-1} w_i^T \quad (17)$$

This matrix is known as the hat matrix for the class i. Finally, the classification process is performed based on the distance criteria formulated using Eq. (18).

$$C^* = \arg\min_i \{ \| v - \tilde{v}_i \| \}, i = 1,2,...,C \quad (18)$$

3.3.3.4. Linear discriminant classifier

Linear discriminant analysis (LDA) is a well-established pattern classification method. For a two-class problem, it determines a projection vector W to maximize the between-class scatter matrix while minimizing the within-class scatter matrix in the feature space [31].

The training process of an LDA classifier is performed through scattering matrix analysis. The within and between-class scatter matrices are calculated, as shown below [32].

$$S_w = \sum_{i-1}^{M} \Pr(C_i) \sum i \quad (19)$$

$$S_b = \sum_{i=1}^{M} \Pr(C_i)(m_i - m_0)(m_i - m_0)^T \quad (20)$$

Here S_w is the Within-class Scatter Matrix showing the average scatter $\sum i$ of the sample vector x of different class C_i around their respective mean m_i [32]:

$$\sum i = E[(x - m_i)(x - m_i)^T \mid C = C_i] \quad (21)$$

Similarly S_b is the Between-class Scatter Matrix, representing the scatter of the conditional mean vectors m_i it is around the overall mean vector m_0.

Various measures exist for estimating the discriminatory power; the commonly used is [32]:

$$\xi(w) = \frac{\| w^T S_w w \|}{\| w^T S_b w \|} \quad (22)$$

Here w is the optimal discrimination projection and can be obtained through solving the generalized eigenvalue problem [32]:

$$S_b w = \lambda_w S_w w \quad (23)$$

Finally, the linear discriminant function for LDA is [32]:

$$d(x) = w^T (x - \frac{m}{2}) \quad (24)$$

3.3.3.5. Deep belief network

Recent advancements in deep learning approaches have been exploited in many fields, including pattern recognition, with the hope of achieving better results in solving many problems compared to traditional approaches. Deep learning approaches are inspired by the deep structure of the human brain. It seeks to enlarge the number of tasks that can be automated and performed with high accuracy using computers [33].

Deep belief network (DBN) is one of the deep learning approaches that proved its effectiveness in many applications such as generating and recognizing images, video sequences, and motion-capture data. In the proposed system, we have designed a DBN based on RBMs to differentiate between authentic images and forged images. The used deep neural network (DNN) is called DBN-DNN. The idea of this network is given in [34]. The architecture of the used DBN-BNN is given in Fig. 4.

Fig. 4 shows that the proposed DBN-DNN consists of eight layers: an input layer, four hidden layers, and an output layer. The size of the input layer is equal to the number of values that exist in the feature vector, while the size of the output layer is two, which is similar to the number of classes in our problem (fire, no fire). The sizes of the hidden layers are 13, 28, 4, and 2, in order. In the training phase, the backpropagation algorithm has been used with a cross-entropy activation function that is defined by Eq. 25, where the batch size is 100, the number of iteration is 100, and the learning rate is 0.1.

$$E = -\sum_i y_i^{'} \log(y_i) \quad (25)$$

Where y_i is the predicted probability distribution of class I, and $y_i^{'}$ is the actual probability for that class.

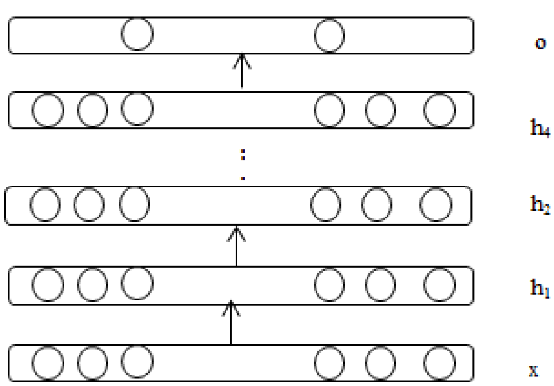

FIGURE 4. The architecture of the used DBN-DNN.

IV. IMPLEMENTATION AND RESULTS ANALYSIS

The proposed system is implemented using Matlab 2018a. The SVM classifier is trained using the cubic kernel function. Several performance measurements are used to evaluate the performance of the different classification models, including:

4. IMPLEMENTATION AND RESULTS ANALYSIS

The proposed system is implemented using Matlab 2018a. The SVM classifier is trained using the cubic kernel function. Several performance measurements are used to evaluate the performance of the different classification models, including:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (26)$$

$$precision = \frac{TP}{TP + FP} \quad (27)$$

$$recall = \frac{TP}{TP + FN} \quad (28)$$

$$Specificity = \frac{TN}{TN + FP} \quad (29)$$

TP is true positive, *TN* is true negative, *FP* is false positive, and *FN* is a false negative. Two experiments have been conducted. In the first experiment, the 10-fold cross-validation approach is used in the evaluation process. In the second experiment, the dataset is divided into two parts: 80% training data and 20% testing data. The DBN classifier has been evaluated only in the second experiment. In the two experiments, the performance of each classification model is evaluated with/without the feature selection step. The results obtained from the first experiment are shown in Table 2.

Table 2: The experimental results of the first experiment

Algorithm	Feature Selection	No. of features	Precision	Recall	Accuracy	Specificity
DT	NONE	17	100%	100%	100%	100%
SVM	NONE	17	99.53%	99.53%	99.59%	99.63%
LR	NONE	17	100.00%	99.53%	99.79%	100%
LD	NONE	17	96.83%	99.53%	98.35%	97.41%
DT	PCA	10	97.62%	95.35%	96.91%	98.15%
SVM	PCA	10	100%	99.53%	99.79%	100%
LR	PCA	10	99.53%	99.53%	99.59%	99.63%
LD	PCA	10	97.72%	99.53%	98.76%	98.15%

Based on Table 2, it is observed that the DT classification model in terms of the used performance metrics without applying the feature selection step. Besides, it is noticed that the LD classifier has the worst performance in terms of the used performance metrics, while the SVM and LR classifiers have a competitive performance when compared to the DT classifier. On the other side, when the feature selection step is performed, it is noticed that the SVM classifier is the best in terms of the used performance metrics, while the DT classifier is the worst. Moreover, the LR and LD come in the second and third ranks, respectively.

Figure 5 represents a comparative evaluation of different performance metrics; precision, recall, accuracy, and specificity using DT, SVM, LR, and LD classifiers. We recognize that DT and LR models have the highest better performance values against SVM and LD classifiers.

Whereas figure 6 describes the same classifiers using exact performance metrics, but we have applied the PCA feature subset selection. The PCA has chosen the top ten features from the available features. In which, we notice that SVM and LR are the better classifiers than DT and LD. In turn, DT has been affected by feature reduction and lost its performance due to feature subset selection, while LD keeps tracks its performance metrics even feature subset selection was applied.

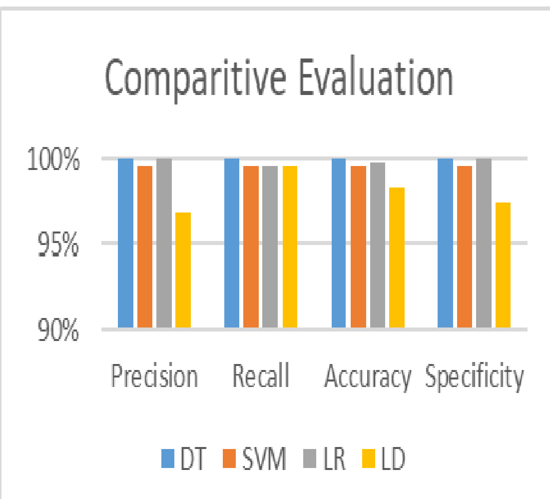

FIGURE 5. Comparative evaluation using four classifiers and 17 features.

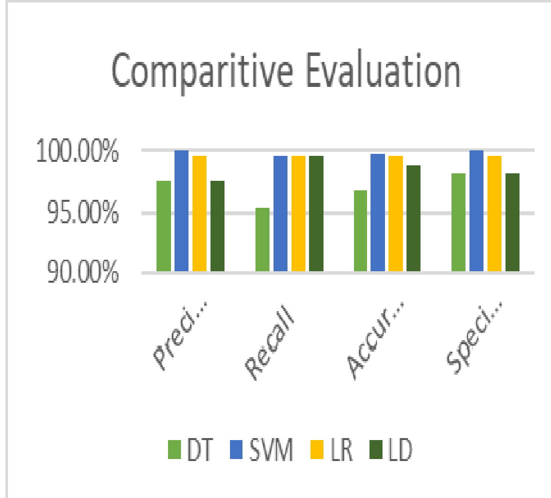

FIGURE 6. Comparative evaluation using four classifiers and ten subset features selected by PCA.

The results obtained from the second experiment are shown in Table 3

Table 3: The experimental results of the second experiment

Algorithm	Feature Selection	No. of features	Precision	Recall	Accuracy	Specificity
DT	NONE	17	97.67%	95.45%	96.94%	98.15%
SVM	NONE	17	99.07%	98.61%	98.97%	99.26%
LR	NONE	17	99.06%	98.59%	98.97%	99.26%
LD	NONE	17	95.96%	99.07%	97.73%	96.65%
DBN	NONE	17	97.65%	96.74%	97.53%	98.15%
DT	PCA	10	95.22%	92.56%	94.64%	96.30%
SVM	PCA	10	94.34%	96.15%	95.88%	95.67%
LR	PCA	10	97.67%	97.67%	97.94%	98.15%
LD	PCA	10	96.33%	96.77%	96.91%	97.01%
DBN	IG	10	97.65%	96.74%	97.53%	98.15%

Based on Table 3, it is noticed that the 10-fold cross-validation approach has achieved better results for all classification models compared to the 20%-holdout approach. Without the feature selection step, it is noticed that the SVM and LR classifiers have the best performance in terms of all performance metrics except the recall metric, where the best results are achieved by the LD classifier. With the feature selection step, it is noticed that the LR classifier gives the performance followed by the LD classifier. Besides, it is noticed that the feature selection using the information gain approach does not affect the DBN classifier.

Figure 7 represents a comparative evaluation of different performance metrics; precision, recall, accuracy, and specificity using DT, SVM, LR, LD, and DBN classifiers. We recognize that

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SVM and LD models have the highest better performance values against DT, LD, and DBN classifiers. These experiments are performed using hold-out validations.

Whereas figure 8 describes the same classifiers using exact performance metrics, but we have applied the PCA feature subset selection. The PCA has chosen the top ten features from the available features. In which, we notice that DBN and LR are the better classifiers than DT, LD, and SVM. In turn, DT and SVM have been affected by feature reduction and lost its performance due to feature subset selection, while LD keeps tracks its performance metrics even feature subset selection was applied.

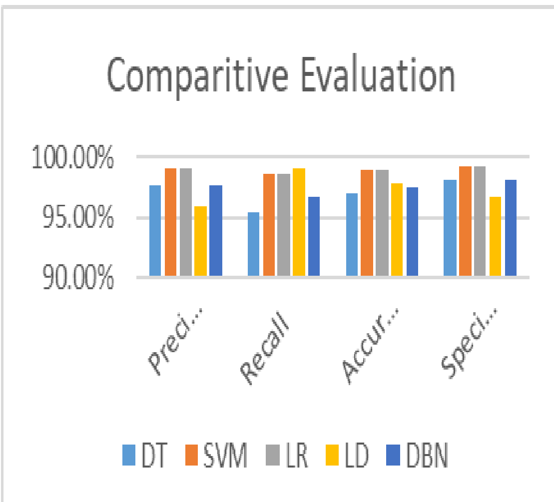

FIGURE 7. Comparative evaluation using four classifiers and 17 features

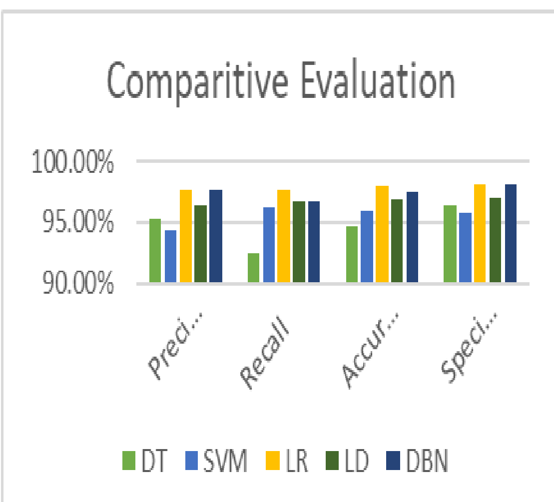

FIGURE 8. Comparative evaluation using four classifiers and ten subset features selected by PCA.

5. CONCLUSION

Early detection of fire ignition can highly improve the fire-fighting process. In this study, we have proposed a fire detection system to detect the fire ignition in the Kuwaiti schools. The proposed system has employed several data mining techniques, whether to perform the feature selection or classification tasks. We used a dataset that has been collected with the help of the General Civil Defense Department in our designed analysis and experiments. The conducted experiment using a 10-fold cross-validation approach has shown that the DT is the best classifier without applying the feature selection. The SVM is the best classifier when the feature selection is performed using the PCA method. In the second experiment where the dataset has been divided 80% training data and 20% testing data, the obtained results have shown that the SVM and the LR are the best classifiers without feature selection process while the LR classifier is the best when the feature selection process. In the future, this work can be extended by using swarm intelligence algorithms to perform the feature selection process, and a more extensive dataset can be used. Another perspective that can be followed is to embed the fire prediction capability into the system.

REFRENCES:

[1] K. Muhammad, J. Ahmad, I. Mehmood, S. Rho, and S.W. Baik, “Convolutional neural networks based on fire detection in surveillance videos,” IEEE Access, vol. 6, pp. 18174-18183, 2018.

[2] Y. C. Li and W. C. Wu, “Visual fire detection based on data mining technique,” in Proc. RVSP, Kaohsiung, Taiwan, 2011, pp. 328-331.

[3] M. Bahrepour, B. J. van der Zwaag, N. Meratnia, and P. Havinga, “Fire data analysis and feature reduction using computational intelligence methods,” In Advances in Intelligent Decision Technologies, Springer, Berlin, Heidelberg, 2010, pp. 289-298.

[4] P. Cortez and A. Morais, “A data mining approach to predict forest fires using meteorological data,” in Proc. EPIA, Guimarães, Portugal, 2007, pp. 512-523.

[5] V. Tsetsos, O. Sekkas, G. Tsoublekas, S. Hadjieythymiades, and E. Zervas, “A forest fire detection system: The Meleager approach,” in Proc. Panhellenic Conference on Informatics, Piraeus, Greece, 2012, pp. 320-326.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[6] N. Alamgir, W. Boles, and V. Chandran, "A model integrating fire prediction and detection for rural-urban interface," in Proc. DICTA, Adelaide, Australia, 2015, pp. 1-8.

[7] W. Schroeder, P. Oliva, L. Giglio, B. Quayle, E. Lorenz, and F.Morelli, "Active fire detection using Landsat-8/OLI data," Remote sensing of environment, vol. 185, pp. 210-220, 2016.

[8] T. Li, M. Ye, F. Pang, H. Wang, and J. Ding, "An efficient fire detection method based on orientation feature," International Journal of Control, Automation, and Systems, vol. 11, no. 5, pp. 1038-1045, 2013.

[9] A. Khatami, S. Mirghasemi, A. Khosravi, and S. Nahavandi, "An efficient hybrid algorithm for fire flame detection," in Proc. IJCNN, Killarney, Ireland, 2015, pp. 1-6.

[10] B. C. Arrue, A. Ollero, and JR M. De Dios, "An intelligent system for false alarm reduction in infrared forest-fire detection," IEEE Intelligent Systems and Their Applications, vol. 15, no. 3, pp. 64-73, 2000.

[11] M. Maksimović, and V. Vujović, "Comparative analysis of data mining techniques applied to wireless sensor network data for fire detection," Journal of Information Technology and Applications, vol. 6, no. 2, Dec, 2013, Accessed on: Oct., 20, 2019 DOI: http://dx.doi.org/10.7251/JIT1302065M, [Online].

[12] K. Angayarkkani, and N. Radhakrishnan, "An intelligent system for effective forest fire detection using spatial data," International Journal of Computer Science and Information Security, vol. 7, no. 1, pp. 202-208, 2010.

[13] M. Maksimovic, V. Vujovic, V. Miloevic, and B. Pericic, "Evaluating the optimal heat detector deployment for fire detection," in Proc. En&T, Russia, pp. 129-134, 2014.

[14] N. Hamadeh, A. Karouni, and B. Daya, "Predicting forest fire hazards using data mining techniques: decision tree and neural networks," In Advanced materials research, vol. 1051, pp. 466-470, 2014.

[15] S. Garcia-Jimenez, A. Jurio, M. Pagola, L. De Miguel, E. Barrenechea, and H. Bustince, "Forest fire detection: A fuzzy system approach based on overlap indices," Applied Soft Computing, vol. 52, pp. 834-842, 2017.

[16] M. Saoudi, A. Bounceur, R. Euler, and T. Kechadi, "Data mining techniques applied to wireless sensor networks for early forest fire detection," in Proc. the International Conference on Internet of things and Cloud Computing, Cambridge, United Kingdom, 2016, pp. 71-77.

[17] H. Lin, X. Liu, X. Wang, and Yunfei Liu, "A fuzzy inference and big data analysis algorithm for the prediction of forest fire based on rechargeable wireless sensor networks," Sustainable Computing: Informatics and Systems, vol. 18, pp. 101-111, 2018.

[18] M. S. Mahmud, M. S. Islam, and M. A. Rahman, "Smart Fire Detection System with Early Notifications Using Machine Learning," International Journal of Computational Intelligence and Applications, vol. 16, no. 2, 2017, Accessed on: Oct, 20, 2019, https://doi.org/10.1142/S1469026817500092, [Online].

[19] S. Al_Janabi, I. Al_Shourbaji, and M. A. Salman, "Assessing the suitability of soft computing approaches for forest fires prediction," Applied computing and informatics, vol. 14, no. 2, pp. 214-224, 2018.

[20] Y. Xie, and M. Peng, "Forest fire forecasting using ensemble learning approaches," Neural Computing and Applications, vol. 31, no. 9, pp. 4541-4550, 2019.

[21] J. S. Wang, W. C. Chiang, Y. L. Hsu, and Y. T. C. Yang, "ECG arrhythmia classification using a probabilistic neural network with a feature reduction method," Neurocomputing, vol. 116, pp. 38-45, 2013.

[22] B. Mwangi, T. S. Tian, and J. C. Soares. "A review of feature reduction techniques in neuroimaging," Neuroinformatics, vol. 12, no. 2, pp. 229-244, 2014.

[23] T. M. Cover, and J. A. Thomas, "Elements of Information Theory," John Wiley and Sons Inc., NY, 1991, pp. 33-36.

[24] K. Koonsanit, C. Jaruskulchai, and A. Eiumnoh, "Band selection for dimension reduction in hyper spectral image using integrated information gain and principal components analysis technique," International Journal of Machine Learning and Computing, vol. 2, no. 3, pp. 248-251, 2012.

[25] X. Wu, , V. Kumar, J. Ross Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan et al., "Top 10 algorithms in data mining," Knowledge and Information Systems, vol. 14, no. 1, pp. 1-137, 2007.

[26] T. Mitchell, "Machine learning," Burr Ridge, IL: McGraw Hill, vol. 45, no. 37, pp. 870-877, 1997.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[27] B. E. Boser, I. M. Guyon, and V. N. Vapnik, “A training algorithm for optimal margin classifiers,” in Proc. of the fifth annual workshop on Computational learning theory, Pittsburgh, Pennsylvania, USA , 1992, pp. 144-152.

[28] A. Mert, N. Kılıç, E. Bilgili, and A. Akan, “Breast cancer detection with reduced feature set,” Computational and mathematical methods in medicine, Vol. 2015, 2015, Accessed on: Oct, 20, 2019, https://doi.org/10.1155/2015/265138, [Online].

[29] X. Pan, X. Zhang, and S. Lyu, “Exposing image forgery with blind noise estimation,” In Proc. the thirteenth ACM multimedia workshop on Multimedia and security, Buffalo, New York, USA, 2011, .pp. 15-20.

[30] M. Koç, and A. Barkana, “Application of Linear Regression Classification to low-dimensional datasets,” Neurocomputing, vol. 131, pp. 331-335, 2014.

[31] S. Luo, E. H. Kim, M. Dighe, and Y. Kim, “Thyroid nodule classification using ultrasound elastography via linear discriminant analysis,” Ultrasonics, vol. 51, no. 4, pp. 425-431, 2011.

[32] W. Zhao, R. Chellappa, and N. Nandhakumar, “Empirical performance analysis of linear discriminant classifiers,” In Proc. CVPR, Santa Barbara, CA, USA, 1998, pp. 164-169.

[33] E. G. Hinton, “Deep belief network”, Accessed on: Oct, 20, 2019, http://www.scholarpedia.org/article/Deep_belief_networks.

[34] G. Hinton. “Deep Belief Nets,” Springer Science Business Media, New York. DOI: 10.1007/978-1-4899-7502-7_67-1, 2014.