AN ONTOLOGY AIDED GSO OPTIMIZED EXTREME LEARNING FOR SITUATION RECOGNITION IN COAL MINING ENVIRONMENT

[1]B Ramesh, [2]Dr. K Panduranga Vittal

[1]Associate Professor, Channabasaveshwara Institute of Technology, Gubbi, Tumkur, Karnataka 572216.

[2]Professor, Department of EEE, National Institute of Technology, Surathkal, Mangaluru, Karnataka 575025.

rbant96@gmail.com

ABSTRACT

A wireless sensor network (WSN) is a wireless network comprising of spatially distributed autonomous devices utilizing sensors for monitoring the physical or environmental situations. WSN has been applied in many fields such as healthcare monitoring, coal mine safety monitoring system and also in military. To detect the seismic activities in the coal mining environment, several techniques such as Bord and Pillar model, Bayesian Decision method etc., were introduced and carried out. In this paper, we have proposed an Ontology aided Fuzzy Cognitive Maps (FCM) based feature correlation extraction technique for the multi attribute sensor data. Further, the Galactic Swarm Optimization (GSO) algorithm optimized Extreme Learning Machine (ELM) is used. The correlation extraction technique gives the better solution to determine the similarity between the semantically related heterogeneous sensor reading data and resolves the semantic ambiguity problem of heterogeneous sensors present in the coal mining Environment.
Keywords: Wireless Sensor Network, Ontology Learning, Fuzzy Cognitive Map, Galactic Swarm Optimization, Extreme Learning Machine

1. INTRODUCTION

The latest advances in wireless communication technologies and micro-electro-mechanical systems have led to great progression in WSNs [1]. WSNs are the wireless networks consist of many spatially distributed sensors with limited processing capability and data gathering to monitor the environmental condition. Because of their ability to monitor and manage situational information for various intelligence services, these networks have become more important [2]. Therefore, the WSNs have been applied in various fields such as Coal Mine Safety Monitoring System, healthcare monitoring and military [3]. To provide support to the working people with special needs in their daily life safety and also to control automation, the mining industries have emerged recently [4].Coal mining industries embody this vision in several ways because, sensors are embedded into everyday environmental activities to collect data which monitors the state of the physical environment and its residents when everyday routines are performed. Then the computational component reasons about the collected information to take an action which optimizes the goals such as safety, productivity and comfort [5], [6].

The coal mining industries are constantly driven by the need to increase the personnel safety, enhance mining productivity, and to achieve environmental sustainability [7], [8]. The underground coal mining contains various hazards, including high energy hydraulic and electrical power systems, personnel proximity to machinery, roof falls and exposure to explosive mine gases and dust [9], [10]. Generally, personnel have been required to work in the hazardous environment and manually control the equipment at close range to ensure the efficient operations of the mining process. The complexity of the manually operating equipment of this scale has led to the full productivity of long wall systems not been achieved. The hazardous or the harsh working environment has potential health and safety concerns for personnel. Sensors in the monitoring environment are embedded anywhere and on any objects or human bodies. They collect the data containing user's motion, location, environmental temperature, activity information, humidity and ambient noise level. Applications which provide customized services to the users are on the basis of this sensor data. However, sensor data exhibits high complexity (huge volumes, different modalities and interdependent relationships between the sources), a dynamism (real-time update and critical ageing), precision, timeliness and accuracy.

A prevalent computing system should not examine itself with the separate pieces of sensor data rather than the information should be interpreted into a higher domain related idea. This higher level concept is a situation, which is an abstract state of affairs interesting for the applications. The ability of using situations lies in their capability to give simple and human readable presentation of sensor data to applications, although shielding applications from the complexities of sensor readings, sensor data noise and inference activities and simultaneously leveraging the structure implicit in the activities are being observed. The system has a significant task of defining and managing these situations. This includes capturing what and how situations are to be recognized from which pieces of contexts, and how different situations are related to each other. Context-aware modeling techniques should be able to understand by the available context sources, their data structure and automatically built internal data models to facilitate them. Additionally, raw context needs to be retrieved and transformed into appropriate context representation models accurately with minimum human intervention. Many popular context modeling techniques are utilized to give a representation of the sensor data. Context models can be static or dynamic. Static models have a predefined set of context information which will be collected and stored. The requirements needed to be taken into consideration when modeling context information are mobility and heterogeneity, dependencies and relationships, imperfection, reasoning, timeliness, efficient context provisioning and usability of modeling formalisms.

Situation recognition is the process of identifying interesting status automatically and changing the physical environments or the entities. With the advancement of context-aware applications in smart spaces, the situation recognition is regarded as a sound method to offer constantly changing situational picture about the observed environment. Nowadays, low-level sensing data obtained from the smart environment (e.g., coal mining environment) is an important source of information for the situation recognition. But, the main problem of sensor-based event recognition is the data obtained from sensors have different degrees of dynamics and uncertainty. This uncertainty arises for several reasons in a sensor network environment such as inaccurate measuring, fault sensors and "dirty" data corrupted by wireless networks owing to the network problems. There are various data-driven techniques such as Machine Learning algorithm and Meta-heuristic optimization aided learning-based techniques which are employed to recognize sensor events in the coal mining environment.

2. RELATED WORKS

Some of the recent research works related to the situation representation and recognition of WSNs in coal mining environment is listed below:

Sudipta Bhattacharjee *et al.* have presented the response time of the developed system for identifying fire in the Bord and Pillar coal mine panel. It uses WSN to observe the exact fire location. In the developed system, the two sensor nodes located at the panel inlet and outlet will detect carbon monoxide and oxygen gas concentration continuously. The output of the two sensor nodes will given as the input to the controller node which is responsible for calculating the difference between the ratio of carbon monoxide and oxygen concentration of the inlet and outlet signal. A threshold limit will be put in the controller node. When the output of the controller node crosses the threshold value, it will first generate an alarm which will be the indication of the occurrence of fire in that working panel. It will result in the activation of the temperature sensor nodes to recognize the exact fire location. The performance of the developed system has been determined via rigorous simulations. The simulation results reveal that the average network delay changes almost linearly with the increase in the number of hops.

Wei Chen *et al.* have analyzed the structural features of the coal mining tunnel on the basis of minimum risk Bayesian decision technique and minimum error Bayesian decision technique. In accordance with the probability of working objects on the choice between the crossways and the tunnels, we propose the prediction approach for selecting the path based on the Bayesian decision, the associated nodes in the branch and tunnel. Based on the conditional probability and the prior probability of the active objects on the tunnel, they can estimate the posterior probability of tunnel selection and also predict the selection probability of the miners on the tunnel in the next moment. This method is the basis for the correlation among the network nodes which can reduce the monitoring nodes, lock the important nodes and improve the accuracy in detecting and tracking the active objects. In decision-making to predict for the selected tunnel of moving target, one method can make the error rate of decision-making minimum. This method can provide a theoretical basis and techniques for monitoring the coal mine safety to prevent and to disclose the hidden hazards of the coal mining environment. Experimental results showed that the Bayesian decision technique can effectively connect the nodes.

Bo Cheng *et al.* have proposed a uniform message space and data distribution model and also a lightweight service mashup approach is implemented. With the help of visualization technology, the graphical user interface of different underground physical sensor devices could be created, which allows the sensors to combine easily with other resources. The proposed mashup middleware is to improve the coal mine monitoring and control automation, which allows the user to create the ad-hoc safety monitoring and automation service intuitively. Our solution includes four phases: 1) to access the sensor data with OSGi-based uniform devices access framework; 2) adopt the publish subscribe mechanism to distribute the sensor data; 3) implement a lightweight services mashup approach that supports the on-the-fly integration of different services to build comprehensive and situational applications; and 4) apply the REST principles to define an extensible interface for end users. The proposed solution is easy to deploy and implement quickly, and may help to improve the coal mine safety monitoring and automation level.

Umar Ibrahim Minhas *et al.* have proposed a WSN-based system for monitoring and event reporting in the underground mine environments. The proposed system which is capable of detecting and identifying events of interest (with 90%success rate) and localization of miners (2–4 m) and roof falls (10–12 m). Moreover, a novel energy-efficient hybrid communication protocol using both periodic and aperiodic modes of communication while adhering to low latency requirement for emergency situations is proposed and implemented. Finally, for intelligent processing of gathered data, a spatio-temporal and attribute-correlated event detection mechanism suitable for the highly unreliable mine environment is described. The system successfully detected and identified the events in all tested cases providing a comprehensive control and monitoring mechanism and tracked location of miners and events required for rescue operations.

Andrzej Janusz *et al.* have investigated and compared the practical approaches for determining the seismic events by using the analytical models constructed on the basis of domain knowledge and the sensor data. For case study, they used a rich data set gathered during the period of five years from various active Polish coal mines. They mainly focused on comparing the prediction quality among the expert methods which serve as a standard in the coal mining industry and state-of-the-art machine learning methods for mining high-dimensional time series data. They described an international data mining challenge organized to facilitate our study and also they demonstrated a technique which we employed to construct an ensemble of regression models able to outperform other approaches used by the participants of the challenge. Finally, they explained how they utilized the data obtained during the competition for the purpose of research on the cold start problem when deploying decision support systems at new mining sites.

Current improvements in pervasive computing and sensor technologies have allowed the contextualized enhancement of business processes capitalizing on the capability to perceive, process, integrate and interpret data of various modalities. Out of the many areas of interest, the representation and recognition of situational events of labors safety is a prominent instance where pervasive computing environments deliver exclusive results for the contextualized monitoring and assessment of behavior or event, such as Coal mining environment. A vital challenge in situation representation is the capability to efficiently combine manifold sources of heterogeneous, noisy and possibly inconsistent data in a way that delivers precise and valuable results. For the representation of heterogeneous sensor data, the knowledge and data-driven approaches are presented in earlier researches. The knowledge-based approaches permit domain knowledge and common sense semantics to be combined into activity models. Data-driven approaches relied on probabilistic and statistical models to represent activities. On the other side recognizing the situations from Coal mining environment various machine learning techniques are utilized like Artificial Neural Networks (ANN), K-Nearest Neighbor (KNN), Support Vector Machine (SVM), and so on. These machine learning algorithms needs to be learned for each individual, it takes maximal iterations to recognize situations form learning environment. Hence, bio-inspired meta-heuristic optimization aided machine learning techniques are provide the way to enhance the situation recognition accuracy with minimal iteration. Therefore, we plan to develop an Ontology aided GSO optimized ELM for situation recognition in coal mining environment.

These areas are reflected in the structure of the paper. After the introduction and overview of related works in Section 2. In Section 3, we present details regarding our proposed approach for the assessment of seismic hazards. This section is followed by Section 4, where we perform an extensive analysis of the results. It aims to answer the question whether it is possible to use traditional prediction models trained on different sensor reading datasets in order to make accurate predictions for a different long wall. Finally, in Section 5, we provide the conclusion of the paper.

3. PROPOSED APPROACH FOR SITUATION RECOGNITION IN COAL MINING ENVIRONMENT

In this paper we develop an Ontology aided GSO optimized EM for situation recognition in coal mining environment. Ontology aided techniques are recently attracted in view of modeling and reasoning over contextual data and environmental activities. Fuzzy Cognitive Map (FCM) is a most dominant technique to model the dynamics of complex system structures and has been utilized in numerous fields, such as engineering, management and so on.

The construction of FCMs has huge significance for its application. Initially for situation representation we constructing a Fuzzy Cognitive Maps aided by ontology matching technique in a holistic manner. The ontology matching technique provide the solution to find the similarity between semantically related heterogeneous sensor reading data and resolves the semantic ambiguity problem of heterogeneous sensors present in coal mining Environment.

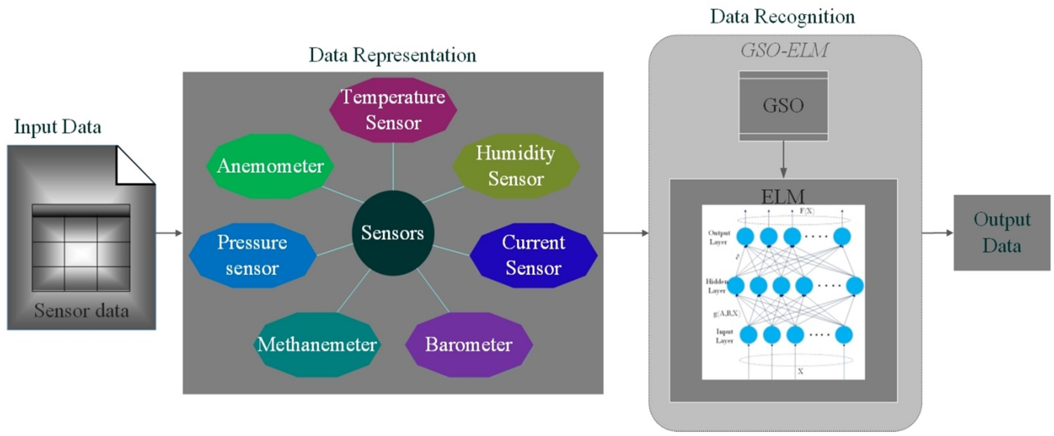

Figure 1: The Schematic Layout of Proposed Approach

This ontology aided FCM technique improves the effective collaboration in the heterogeneous sensors present in coal mining environment. The proposed holistic FCMs also permit to draw extra observations about the underlying heterogeneous sensors, which are not obtainable through the separate FCMs. Finally based on the ontology aided FCM construction, the GSO optimized ELM is employed to recognize situations in SH environment activities.

3.1 Ontology representation for multi attribute sensor data

Nowadays, it is a popular technique to explain context information by means that of ontology modeling, which might depict the context information simply and correctly. Ontology could be a structured system of ideas that covers a particular field and presents the reality in a type of a model. It is also termed as a group of ideas and classes in an area or domain which shows their properties and also the relations between them.

A Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) is a standardized higher ontology that is especially utilized to describe the common objects within the real world. DOLCE Ultra Lite (DUL) is the abridged version of DOLCE. By ignoring the complicated factors such as the relation network, time index and so on, it tries to lower the problem of modeling information concept ontology. Semantic Sensor Network (SSN) ontology is developed by the W3C Semantic Sensor Network Incubator Group. Its final version was shaped in 2011. It provides a good framework to explain the context data by reusing the DUL ontology.

But, SSN Ontology doesn't give the outline of the domain concept, location, time, etc., however the description of sensors, related concepts and sensor observations. Therefore, the application in some specific areas must be further more expanded on the basis of upper ontology. By means of the inheritance of SSN Ontology and DOLCE Ultra Lite, combined with the domain characteristics and actual demand of coal mine, the Coal Mines Semantic Sensor Network Ontology (CMSSN) is constructed to inherit some classes similarly some properties from the DUL and SSN Ontology.

Meanwhile, consistent with the characteristics of the coal mine domain, to fulfill the fundamental desires of service system, some specialized classes and a few attributes are generated in the CMSSN. On one hand, some classes in CMSSN are inherited directly from DUL and SSN; on the other hand, the derived classes are analyzed and reused in CMSSN Ontology.

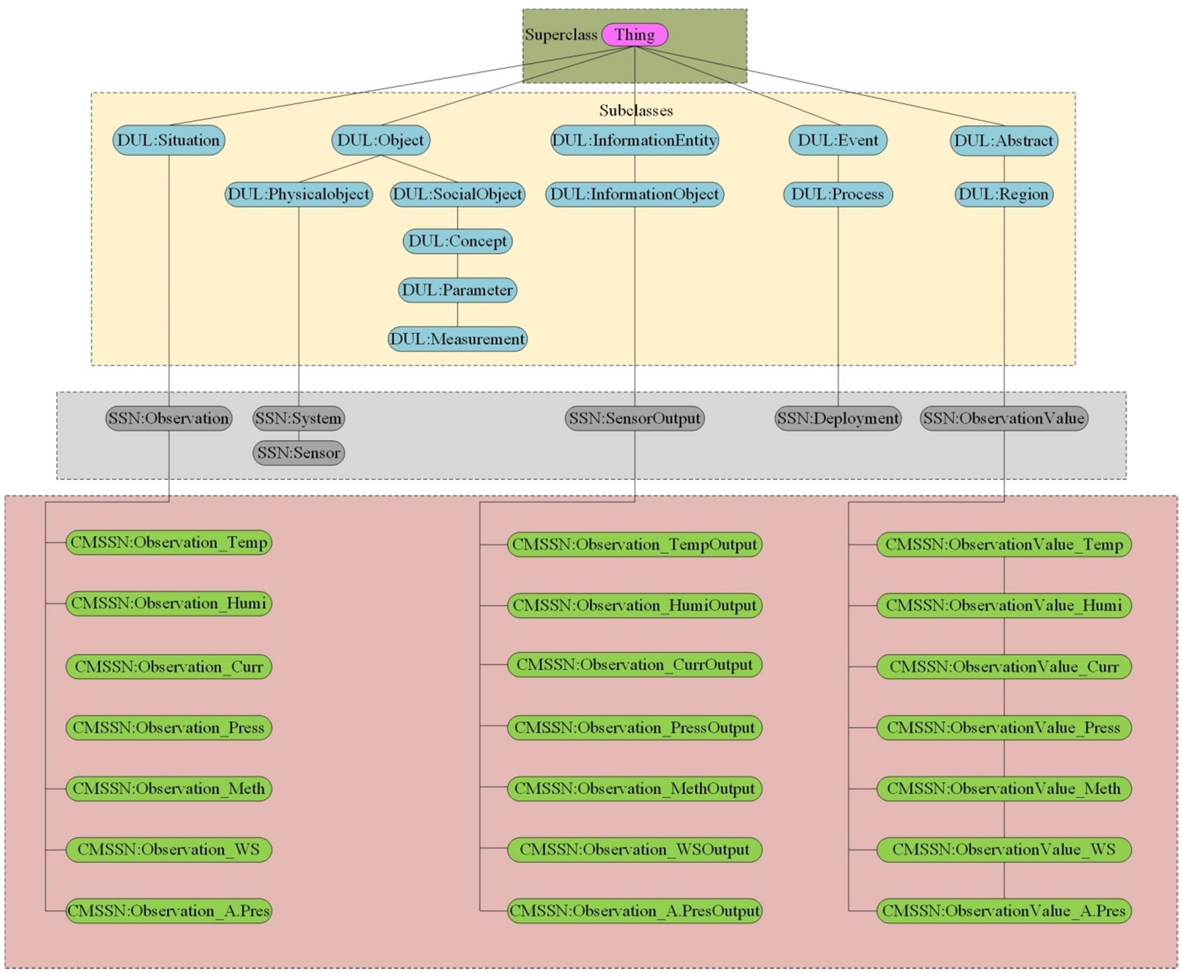

Figure 2: Ontology Representation

3.2 FCM Representation of Contest the Sensor DataStream in Coal Mining Environment

A Fuzzy Cognitive Map (FCM) is a directed graph which includes the factors such as nodes or concepts and edges. The events and elements are called as nodes and the causal relationship labeled with fuzzy values are called as edges. It shows the causal relationship among the concepts. It can be represented as a graphical representation of the learning about the given system. It is a combination of both fuzzy logic and cognitive mapping. Cognitive mapping is on the basis of graph theory as well as indices and calculations. The influence of the factors on each other is iteratively calculated with the help of the originated method in the neural network approach. The results present the trends within the system, once the network has stabilized. FCM also provides the possibility to process simulations and to determine the outcomes of possible scenarios.

In the proposed method, sensors are used to detect the behavior of the environment. The sensors are detecting the parameters such as temperature, vibration etc. The sensor readings are stored in the server/controller continuously. The controller node has the threshold value for each parameter. If the sensor reading exceeds the threshold value, the controller will give the alarm signal. For representing the sensor readings, Fuzzy Cognitive Map technique is used.

Usually, there are two types of functions are used in the FCM framework. One is unipolar sigmoid function and that is represented as,

$$F(x) = \frac{1}{1+e^{-\alpha x}}. \tag{1}$$

Where, $\alpha > 0$ calculates the steepness of the function F and transforms the content of the function in the interval [0, 1].

Another threshold function which is used and transforms the content of the function in the interval [1, 1] is

$$F(x) = \tan H(x). \tag{2}$$

Based on the method, the function is selected. The overall mathematical expression of FCM is

$$V^{t} = F\left(V^{t-1} w_0 + K_2^j V^{t-1}\right) \tag{3}$$

Where, V is the state vector and w_0 is the weight matrix.

The list of correlation parameters used for feature extraction

Max -The maximum value of the readings in the window

Min -The minimum value of the readings in the window

MaxMinDiff -The difference between the max and min mean the average value of readings in a window

PercentileX - *X*th percentiles for the readings, where:

$X \in \{2, 5, 10, 15, 20, 25, 30, 50, 70, 75, 80, 85, 90, 95, 98\}$

Percentiles5Diff the subtraction of percentiles 95% and 5%

StdDev - Standard deviation of the readings

Kurtosis -The measure of the distribution tail extremity

Skewness-The measure of the distribution asymmetry

3.3 Extreme Learning Machine (ELM) for recognition

Extreme Learning Machine is a feed-forward neural network for the regression, sparse approximation, feature learning, compression, clustering and for classification with a single or multi layers of hidden nodes in which the parameters are randomly generated and they are need not be changed. The hidden nodes are assigned randomly and are never updated. But they can be inherited from their ancestors. In many cases the output weights of hidden nodes are generally learned in a single layer which is essential for learning a linear model.

First, the ELM was proposed for the single hidden layer feed-forward neural network (SLFN), which was then changed to single hidden layer feed-forward network, since the hidden nodes need not be like a neuron. The hidden nodes of the ELM are not only independent of the training dataset but also they are independent to each other. Before seeing the training data, the ELM can create the hidden node parameters.

The output function of SLFN is represented as $F_H(X)$.

$$F_H(X) = \sum_{i=1}^{H} \partial_i G_i(X) = \sum_{i=1}^{L} \partial_i g(A_i, B_i, X), X \in r^d, \partial_i \in r^m \tag{4}$$

Where, H represents the number of hidden nodes, G_i denotes the activation function of i[th]hidden node, (A_i, B_i) are the parameters of the hidden layer. The number of input nodes is represented as d and the number of output nodes as m. ∂_i is the output weight which connects the hidden nodes and the output nodes.

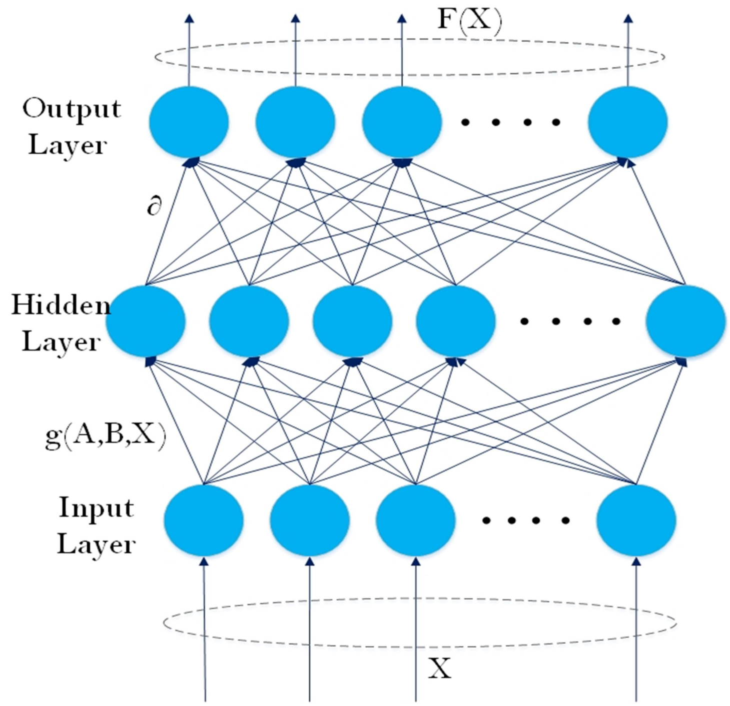

Figure 3: Structural diagram of SLFN

The activation function for additive hidden nodes can be denoted as follows.

$$G_i = g(A_i, B_i, X) = G(A_i X + B_i), A_i \in r^d, B_i \in r \quad (5)$$

For Radial Basis Function, the activation function can be represented as

$$G_i = g(A_i, B_i, X) = G(B_i \parallel X - A_i \parallel), A_i \in r^d, B_i \in r^+ \quad (6)$$

If the training set contains n number of samples $\{(X_i, T_i)\}_{i=1}^{n}$ where X_i can be represented as $X_i = [X_{i1}, X_{i2}, X_{i3}, \ldots\ldots X_{id}]^P \in r^d$ and T_i can be formulated as $T_i = [T_{i1}, T_{i2}, T_{i3}, \ldots\ldots, T_{im}]^P \in r^m$, then the SLFN can approximate the n samples with zero error and that can be denoted as follows.

$$\sum_{j=1}^{n} \parallel Y_j - T_j \parallel = 0 \quad (7)$$

Where, Y is the actual output of the SLFN.

Matrix form of the eqn, $\sum_{i=1}^{H} \partial_i g(A_i, B_i, X_j) = T_j, j = 1,2,3,\ldots., n$ is $Q\partial = P$.

$$Q = \begin{bmatrix} q(X_1) \\ q(X_2) \\ q(X_3) \\ . \\ . \\ . \\ q(X_n) \end{bmatrix} = \begin{bmatrix} g(A_1, B_1, X_1).....g(A_H, B_H, X_1) \\ g(A_1, B_1, X_2).....g(A_H, B_H, X_2) \\ g(A_1, B_1, X_3).....g(A_H, B_H, X_3) \\ . \\ . \\ . \\ g(A_1, B_1, X_n).....g(A_H, B_H, X_n) \end{bmatrix}_{n \times H} \quad (8)$$

$$\partial = \begin{bmatrix} \partial_1^p \\ \partial_2^p \\ \partial_3^p \\ . \\ . \\ . \\ \partial_H^p \end{bmatrix}_{H \times m}, \quad P = \begin{bmatrix} T_1^P \\ T_2^P \\ T_3^P \\ . \\ . \\ . \\ T_n^P \end{bmatrix}_{n \times m} \quad (9)$$

The main feature of ELM is the parameters of the hidden layer are randomly created and kept fixedly without any change by the iteration. An important step in ELM is to find the output weights ∂_i. It is solved by reducing the training error and also the rule of the output weights.

3.4 Galactic Swarm Optimization (GSO) Algorithm

The original PSO algorithm was developed by Kennedy and Eberhart. It was encouraged by swarming like schooling behavior in fish and flocking behavior in birds. Both the swarms do not process any global optimization but they provide a technique for confusing predators and predator avoidance. The PSO algorithm is related to swarms that do not perform global optimization itself, since it is stimulated by swarming behavior. It is not required for other PSO inspired heuristics besides they perform well on benchmark test functions.

The GSO algorithm follows the movement of galaxies, stars and super clusters of galaxies in the cosmos. Stars are not dispensed uniformly in the cosmos but clustered into galaxies which are not uniformly distributed. On a large scale the individual galaxies appear as point masses. In the GSO algorithm, the attractiveness of stars inside a galaxy to large masses and galaxies themselves to other large masses is imitated as follows.

Initially, according to the PSO approach individuals in each subpopulation are attracted to better solutions in the subpopulation. Then, each subpopulation is assumed to be showed by the best solution found by the subpopulation and that can be treated as a super-swarm. The individual in the super-swarm including the best solutions found by each subpopulation also move according to the PSO model.

This technique is normal and the operations of the swarm and the super-swarm can be achieved by using other population based optimization systems. Being attracted towards the global best, this is analogous to the members of a particular swarm. On a larger scale, individual galaxies appear to be point masses and that can be clustered with nearby galaxies to form super clusters of galaxies. In the GSO algorithm, a cluster of galaxies is analogous to the super-swarm and a galaxy of stars is analogous to the sub-swarm. With the Centre of Mass (CM) of the galaxies the cluster of galaxies is identified.

Each individual in the super-swarm shows the global best solution of individual sub-swarms. This analogy of stars is clustered into galaxies, the galaxies are clustered into super clusters and that can be extended to further concepts such as super clusters are introduced. In the proposed system, the analogy is limited to clusters of galaxies and galaxies of stars. In the GSO model, the swarm is a set E of D−tuples comprising elements ($e_j^{(i)} \in \rho^D$) consists of N partitions, called sub-swarms Ei and the size of each of them is M. Randomly, the elements of E are initialized within the search space $[e\min, e\max]^D$. The entire swarm framework is defined by

$$
\begin{aligned}
&E_i \subset E : i = 1,2,3,\ldots,N \\
&e_j^{(i)} \in E_i : j = 1,2,3,\ldots,M \\
&E_i \cap E_j = \phi : if\ i \neq j \\
&W_{i=1}^{N} E_i = E
\end{aligned}
\tag{10}
$$

Where E_i is the swarm of size M, the velocity of each element is denoted as $V_j(i)$ and the personal best is represented as $P_{bj}(i)$.

Strict subdivision of search region is not imposed in this approach. The motion of a subswarm in E_i is independent and has no influence on another subswarm E_j for $i \neq j$, by means of that enabling comprehensive and an unaffected search is possible. A galactic best is defined as G_b, to take an advantage of the newly found exploration ability of multiple subswarms. Whenever any of the global bests $G_b(i)$ assumes a lower function value $F(G_b(i)) < F(G_b)$, it is necessary to be updated. By updating G_b, this method keeps the record of its best solution. When multiple swarms are present, a synergistic effect can be observed rather than having a single swarm exploring towards a particular direction, resulting in a much improved exploration. All the sub swarms independently explore the search space freely on its own. The iteration starts by determining the position and the velocity. The position and the velocity updates are expressed as

$$V_j(i) \leftarrow I_{w1}V(i) + C_1R_1\left(P_{bj}(i) - e_j^{(i)}\right) + C_2R_2\left(G_b(i) - e_j^{(i)}\right) \quad (11)$$

$$e_j^{(i)} \leftarrow e_j^{(i)} + V_j(i)` \quad (12)$$

$$I_{w1} = 1 - \frac{m}{s1+1} \quad (13)$$

$$R_i = W(-1,1) \quad (14)$$

Where I_{w1} is the inertial weight and R_1, R_2 are the random numbers, m is the current integer iteration number and it varies from o to s1. R_i varies from -1 to 1.

In the next stage, the global bests are clustered to form the super clusters. By collecting the global bests from sub swarms E_i, a new super swarm Z is generated.

$$z^{(i)} \in Z : i = 1,2,3,\ldots,N \quad (15)$$

$$z^{(i)} = G_b(i) \quad (16)$$

The position vector and the velocity are expressed as follows.

$$V_j(i) \leftarrow I_{w2}V(i) + C_3R_3\left(P_b(i) - z^{(i)}\right) + C_4R_4\left(G_b - z^{(i)}\right) \quad (17)$$

$$z^{(i)} \leftarrow z^{(i)} + V(i) \quad (18)$$

It can improve the exploitation because the super swarm concentrates on the global bests from the subs warm. The super swarm uses the best solution which is already determined by the sub swarms and therefore exploits the computed information. The super swarm does not process independent exploration, even though the super swarm individuals are more spread out when compared to the individuals of the subs warm.

3.5 Optimized ELM Learning for Situation recognition

To develop an ensemble system, the steps incorporate how to choose the optimal base classifiers and select the ensemble weights for the chosen base classifiers. Here, double optimization is performed by GSO to choose the base ELM and to optimize the ensemble weights comparing to every selected ELM. Concerning choosing the base classifiers, we utilize GSO to look through the optimal ELM sets by considering both the classification performance and the diversity of the ensemble framework, which influences the ensemble system to hike high convergence ability with similarly more diversity.

Besides, to minimize the complexity of the ensemble model without impacting the convergence exactness of the ensemble system, a few of base ELM with much lower ensemble weights than others is pruned from the ensemble framework. Since the proposed ensemble is based on ELM with double optimization on the basis of GSO. The GSO constructs the effective ensemble system with two stages. In the primary stage, GSO is utilized to choose the optimal base ELM and in the secondary stage GSO is used to select the ensemble weights.

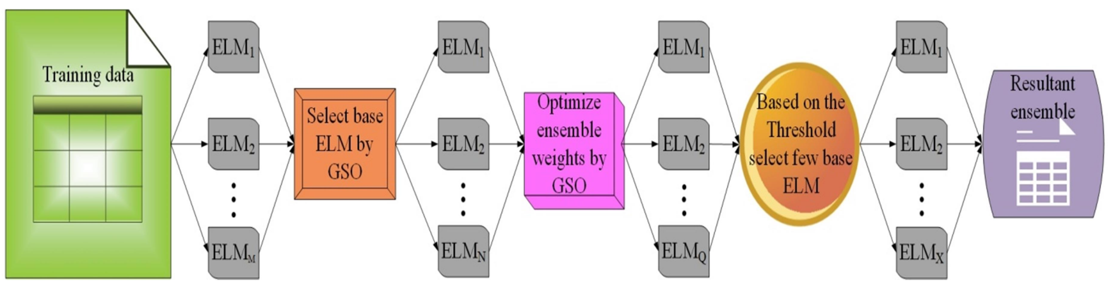

Figure 4: GSO based ELM

The steps required for the model are explained as follows:

A) The dataset is partitioned into two. One is training data and the other is testing data. Using the training dataset corresponding ELMs are generated.

B) Initialize the position and the velocity of each particle randomly. By using the GSO algorithm select the base ELM.

C) Calculate the fitness function of the ith particle by using the below expression.

$$f(X_i) = \sum_{k}^{N_s} \left(D_k \bullet A_k^i\right) \Big/ N_s$$

where D_k is the desired output, A_k^i is the actual output of the ith ensemble system and N_s represents the number of samples.

D) Using the fitness function, update the personal best P_b and the global best G_b values of all the particles.

$$P_b = \begin{cases} X_i & \left((f(X_i) - f(P_b)) > \alpha\right) or (| f(X_i) - f(P_b) | < \alpha \; and \; div(X_i) > div(P_b)) \\ P_b & else \end{cases}$$

$$G_b = \begin{cases} X_i & \left((f(X_i) - f(G_b)) > \alpha\right) or (| f(X_i) - f(G_b) | < \alpha \, and \, div(X_i) > div(G_b)) \\ G_b & else \end{cases}$$

where $div(X_i)$ is the diversity of the ensemble system.

$$div(X_i) = \sqrt{2 \sum_{k=1}^{N_i=1} \sum_{l=k+1}^{N_i} \left(\| IW_k - IW_l \|_2^2 + \| HV_k - HV_l \|_2^2 + \| OW_k - OW_l \|_2^2 + \| A_k - A_l \|_2^2 \right) / N_i \times (N_i - 1)}$$

where IW is the input weight matrix, HV represents the hidden vector, OW denotes the output weight matrix and N_i is the number of base ELM.

E) In the second stage, GSO will optimize the ensemble weight of the base ELMs obtained from the first stage. The optimization process will be same as the first stage but some more details should be verified.

F) The same process will be repeated and the personal best and the global best are updated in the second stage. Redundant base ELMs are deleted and the optimal ensemble is obtained.

The proposed technique will be implemented in MATLAB platform and the performance is evaluated and compared with existing machine learning techniques. The proposed technique will includes the following steps. Initially for situation representation we constructing a Fuzzy Cognitive Maps aided by ontology matching technique in a holistic manner. The ontology matching technique provides the solution to find the similarity between semantically related heterogeneous sensors present in coal mining environments. Finally based on the ontology aided FCM construction, the Galactic Swarm Optimization (GSO) algorithm optimized ELM (Extreme Learning Machine) is employed to recognize situations in Coal mining environment activities. The proposed technique will be validated in terms of accuracy, precision, recall and compared with earlier machine learning techniques.

4. RESULTS AND DISCUSSION

In this paper, not only do we examine the consistency of expert approaches on a data set that covers a period of over five years of seismic readings recorded at different coal miners working places, but we also compare it to results of many other prediction models.

4.1 Dataset description

AAIA'16 Data Mining Challenge: Predicting Dangerous Seismic Events in Active Coal Mines took place between two years period. It was provided by Knowledge Pit platform, under auspices of 11th International Symposium on Advances in Artificial Intelligence and Applications which is a part of the FedCSIS conference series.

The dataset was related to the assessment of safety conditions in underground coal mines with regard to seismic activity and early detection of seismic hazards. In particular, the data set was composed of readings from sensors monitor the seismic activity perceived at long wallsof different coal mines and measure energy released by, so called, seismic bumps. Each case in the data was described by a series of hourly aggregated sensor readings from a 24 h period.

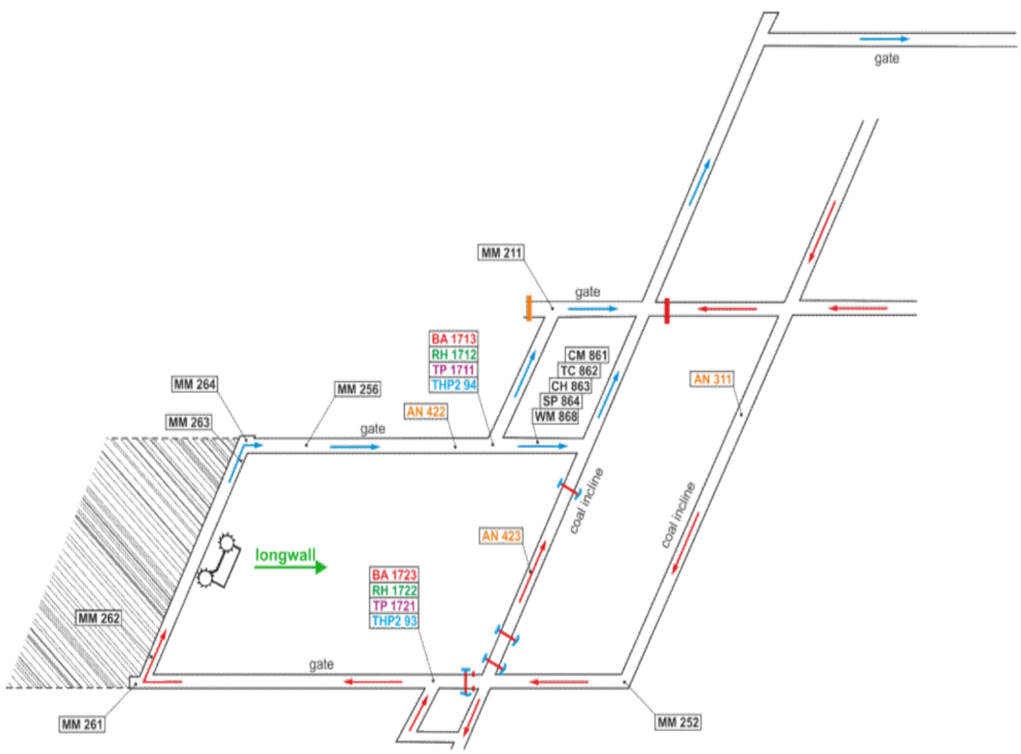

Figure 5: Layout of coal mining process scheme

4.2 Performance Evaluation

In this section, we discuss the experimental results on AAIA'16 data mining challenge dataset contains multi attribute sensor reading values related to underground coal mining environment. Initially, the multi attribute sensor reading data's are represented using the help of ontology language, further fuzzy cognitive map based method is employed to discover the correlation of multi attribute sensor reading data for efficient statistical feature extraction. To evaluate the efficacy of proposed approach we compare traditional prediction model such as SVM (Support Vector Machine), KNN (K-Nearest Neighbor) and Artificial Neural Networks (ANN). However, the best recognition score was obtained using the proposed GSO optimized ELM classifier.

Table 1: Performance Evaluation of GSO-ELM training and testing

Dataset	Training Accuracy	Training Time(sec)	Testing Accuracy	Testing Time(Sec)
D1	0.9040	1.0822	0.9456	0.4630
D2	0.8988	1.5553	0.9405	0.4096
D3	0.8969	1.3661	0.9385	0.4696
D4	0.9064	0.9325	0.9480	0.4479
D5	0.9350	0.9186	0.9766	0.3919

The table 1 depict the performance evaluation of proposed GSO-ELM prediction models in terms of training and testing accuracy as well as time of execution for different sensor reading dataset in coal mining environment.

The following figure 3,4,5,6 shows the recognition performance of proposed approach with traditional prection model such as SVM, KNN and ANN. From the figures validation we observe our proposed approach is provides betters outcomes in terms of accuracy, f1-score, precision and recall.

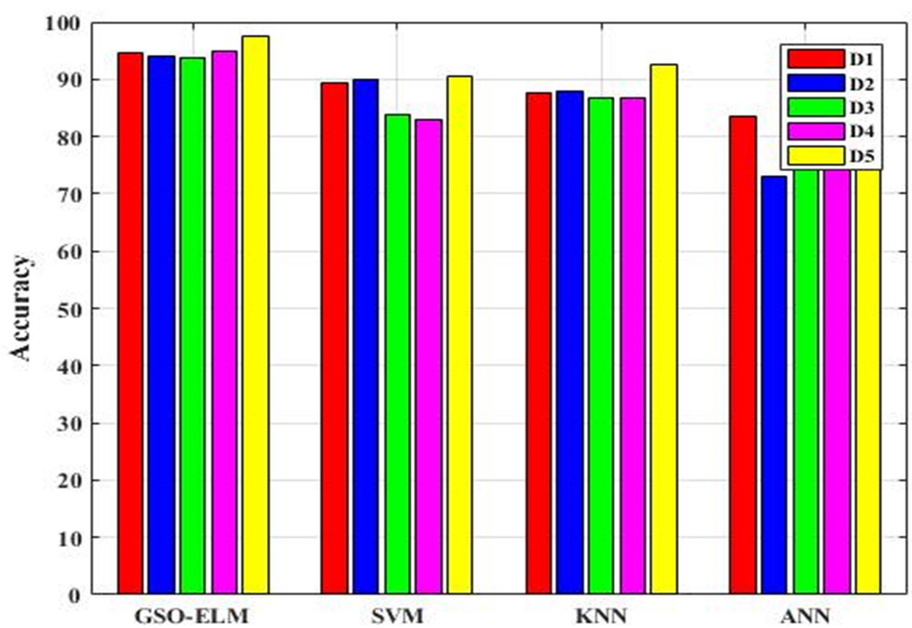

Figure 6: Accuracy Comparison of prediction models

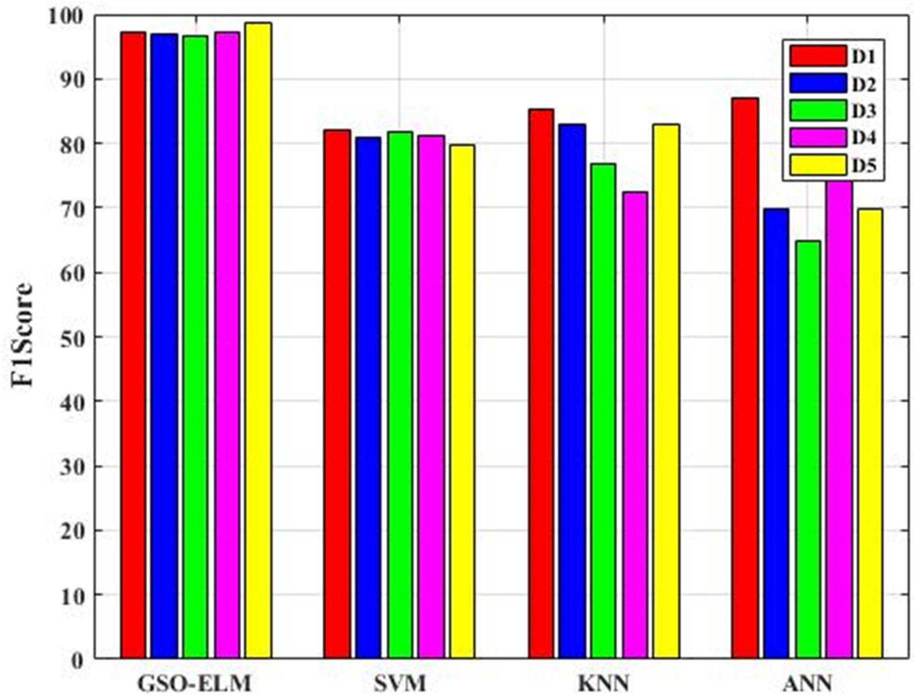

Figure 7: F1 Score Comparison of prediction models

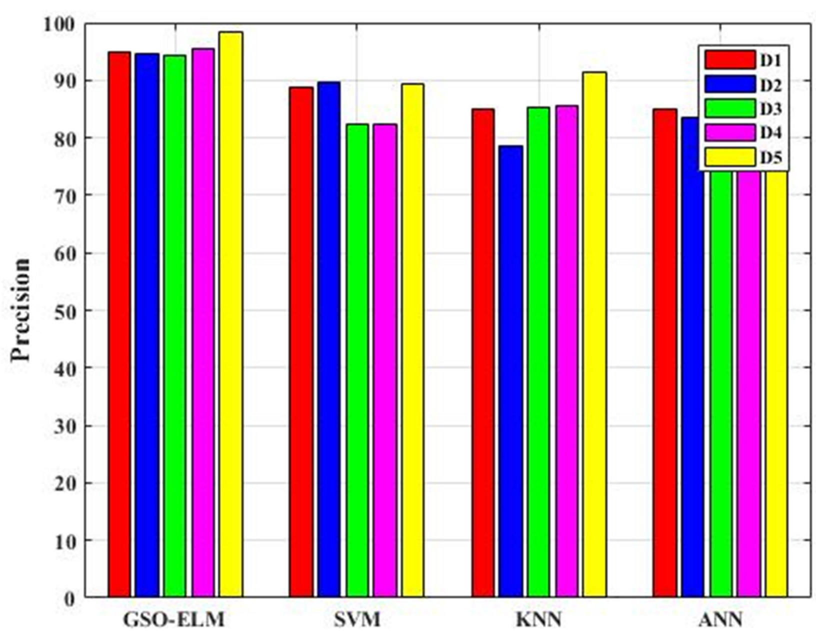

Figure 8: Precision Comparison of prediction models

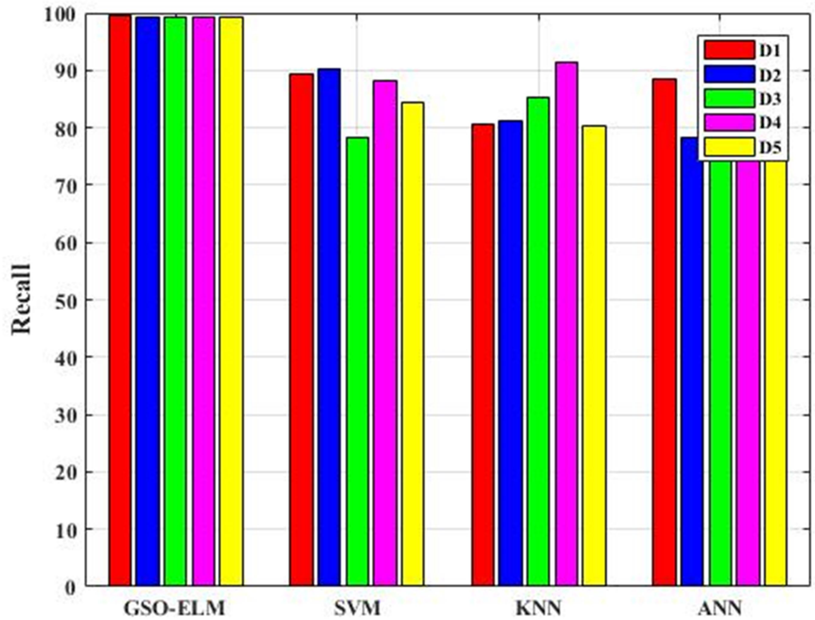

Figure 9: Precision Comparison of prediction models

A comparison of the selected solutions to predictions that were based solely on assessments made by experts revealed that more complex models were able to quickly attain significantly higher scores for working sites with available training data. In the case of the remaining working sites the advantage of complex prediction models was not that clear.

The average results for selected models were only slightly higher, however, for a part of the investigated solutions, the difference was much more favorable than for others. This observation makes it possible to formulate a general strategy for dealing with the cold start problem: in the case of new working sites, start predicting seismic hazards using the expert methods and concurrently gather data for training a more sophisticated prediction algorithm. Initiate your model using data from other working sites and then adjust it using the newly obtained data. Periodically compare the performance of your model to results of the expert assessments. Carefully tune the threshold for issuing the warnings on scores predicted by your model and switch to your predictions when they become more accurate

On the same training and test sets, we run the recognition experiment using KNN, SVM, ANN and ELM. ELM outperformed ANN in terms of accuracy (see above Table 1). However, we observed only a small difference in accuracy between the ELM and SVM classifiers. In terms of training time, the ELM outperformed SVM, KNN and ANN by a huge margin (Table 3). As our eventual goal is to develop a real-time multimodal sentiment analysis engine, so we preferred the ELM as a classifier which provided the best performance in terms of both accuracy and training time.

5. CONCLUSION

The outcome of the proposed approach permit to firmly gives feasible way to develop an intelligent information system which can reveal early warnings about seismic hazards in underground coal mining environment at least as effective as the earlier machine learning Moreover, the proposed approach results demonstrated on coal mining challenge dataset, the prediction performance of proposed approach is improved when compared to earlier methods such as SVM, KNN and ANN in terms of Accuracy, F1 Score, Precision and Recall. However, the valuations proposed by domain experts are still very beneficial and should not be neglected. The statistical features to be an significant information source for the machine learning algorithms to construct intelligent warning system, but they can also be used throughout anew mining sites model deployment period. Finally, the conducted experiments show that the excellence of predictions can be better by merging data obtained from sensors with assessments proposed by domain experts.

REFERENCES

1. Baronti P, Pillai P, Chook VW, Chessa S, Gotta A, Hu YF, “Wireless sensor networks: A survey on the state of the art and the 802.15. 4 and ZigBee standards”, Computer communications, Vol.30, No.7, pp.1655-1695, 2007.
2. Yick J, Mukherjee B, Ghosal D, “Wireless sensor network survey”, Journal of Computer networks”, Vol.52, No.12, pp.2292-2330.
3. Alemdar H, Ersoy C, “Wireless sensor networks for healthcare: A survey”, Journal of Computer Networks, Vol.54, No.15, pp.2688-2710.
4. Islam SR, Uddin MN, KwakKS,“TheIoT: Exciting Possibilities for Bettering Lives: Special application scenarios”, IEEE Consumer Electronics Magazine, 2016 Apr; Vol.5, No.2, pp.49-57, 2016.
5. Acero AR, Cano AM, Builes JA,“SCADA System for Detection of Explosive Atmospheres in Underground Coal Mines through Wireless Sensor Network”,IEEE Latin America Transactions,Vol.12,No.8,pp.1398-1403,2014.

6. Li M, Liu Y, Chen L,"Nonthreshold-based event detection for 3D environment monitoring in sensor networks", IEEE Transactions on Knowledge and Data Engineering, Vol.20,No.12,pp.1699-1711,2008.
7. Singh RD,"Principles and practices of modern coal mining". New Delhi: New Age International, 2005.
8. Van Duin S, Meers L, Donnelly P, Oxley I, " Automated bolting and meshing on a continuous miner for roadway development", International Journal of Mining Science and Technology,Vol.23,No.1,pp.55-61,2013.
9. Weeks JL,"Health hazards of mining and quarrying. Geneva: International Labour Organization", 2011.
10. Ralston J, Reid D, Hargrave C, HainsworthD,"Sensing for advancing mining automation capability: A review of underground automation technology development", International Journal of Mining Science and Technology, Vol.24, No.3, pp.305-310, 2014.