t-SNE and variational auto-encoder with a bi-LSTM neural network-based model for prediction of gas concentration in a sealed-off area of underground coal mines

Prasanjit Dey1 · K. Saurabh1 · C. Kumar1 · D. Pandit1 · S. K. Chaulya1 · S. K. Ray1 · G. M. Prasad1 · S. K. Mandal1

Accepted: 10 September 2021 / Published online: 5 October 2021 © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

A deep learning network is introduced to predict concentrations of gases in the underground coal mine enclosed region using various IoT-enabled gas sensors installed in a metallic gas chamber. The air is sucked automatically at specific intervals from the sealed-off site utilizing a solenoid valve, suction pump, and programmed microprocessor. The gas sensors monitor the gas content in the underground coal mine and communicate gas concentration to the surface server room through a wireless network and cloud storage media. The t-SNE_VAE_bi-LSTM model is proposed in this study as a prediction model that combines the t-SNE, VAE, and bi-LSTM networks. The proposed model’s t-SNE method aims to minimize the dimensionality of the recorded gas concentration; and VAE layer intends to retrieve the inner characteristics of low-dimensional gas concentration. Finally, the given model’s Bi-LSTM layer tries to forecast the concentrations of CH₄, CO₂, CO, O₂, and H₂ gases. The proposed model’s prediction accuracy is compared with the existing two models, namely auto-regressive integrated average moving (ARIMA) and chaos time series (CHAOS). The experiment findings demonstrate that the t-SNE_VAE_bi-LSTM model forecasted mean square error (MSE) is more accurate, and it has lesser MSE value of 0.029 and 0.069 for CH₄; 0.037 and 0.019 for CO₂; 0.092 and 0.92 for CO; 1.881 and 1.892 for O₂; and 1.235 and 1.200 for H₂ than the ARIMA and CHAOS models, respectively.

Keywords Variational auto-encoder · Bi-LSTM · t-SNE · IoT · Gas concentration · Sealed-off area · Underground coal mine

1 Introduction

Due to the harsh and unpredicted working conditions in the underground coal mine, many accidents occurred (Qiao and Zeng 2011; Wang et al. 2013). Furthermore, as per Mine Safety and Health Administration (MSHA), the explosion of flammable gases was occurred in underground mines (MSHA 2018). As a result, it had reduced the production in the affected mines. Therefore, there is a necessity for designing a system that continually monitors and forecasts gas concentrations in underground coal mine enclosed regions for the safety of workers and the early opening of the enclosed site to begin production. Some traditional sensing and monitoring technologies are available for monitoring gas concentrations (Kumar et al. 2013; Mandal et al. 2013; Chaulya and Prasad 2016). However, these traditional methods are incapable of processing massive amounts of multidimensional data provided by various sensors.

In this study, the IoT-enabled gas sensors (CH₄, CO₂, CO, O₂, and H₂) are deployed in a gas collecting chamber installed outside the enclosed site wall of an underground coal mine. The inlet of the metallic gas chamber is connected to a solenoid valve followed by a metallic pipe inserted deep into the fireside through a sealed-off wall. The solenoid valve automatically opens in the fixed time interval for collecting different gas concentrations of a sealed-off area. The five gas sensors gather the appropriate gas concentrations and transmit them to the prediction system via a wireless local area network (WLAN). After
collecting 2-min sampling, the solenoid valve closes automatically and opens after a predefined programmable interval. The multidimensional data are generated from IoT-enabled gas sensors. Then generated data send to the prediction model, where the model is trained. After completing the training and validation processes, the prediction model predicts the respective gas concentration of a sealed-off area and sends the prediction result to the cloud storage. Finally, the mine management accesses the cloud storage from the surface control room and monitors the sealed-off area’s environmental condition.

Recently, Huang and Kuo (2018) have developed a deep convolutional neural network (CNN)-long short-term memory (LSTM) method for weather forecasting in a smart city. On the other hand, deep learning networks have been broadly used in natural language processing (NLP), computer vision, and object recognition (He et al. 2016; Collobert and Weston 2008). However, deep learning networks effectively process complex multidimensional data (Rashid and Rehmani 2016; Jo and Khan 2018; Saeed et al. 2019). But this method does not correlate the gas concentration, air velocity, and temperature parameters of the underground mine. Hence, this method cannot accurately predict the concentrations of gases in the enclosed region. Therefore, the combination of a deep learning network and IoT technologies has been adopted to monitor the underground mine environment (Rashid and Rehmani 2016; Jo and Khan 2018; Muduli et al. 2019; Saeed et al. 2019).

To achieve this goal, a reliable and efficient deep learning model has been developed that can predict the error-free gas concentration in real time for an underground mine’s sealed-off area. Thus, a deep learning method is proposed as a prediction model to combine t-distributed stochastic neighbor embedding (t-SNE), variational autoencoder (VAE), and bidirectional LSTM (bi-LSTM) network, which is named as a t-SNE_VAE_bi-LSTM model. The t-SNE algorithm preprocesses multidimensional time series data from the CH4, CO2, CO, O2, and H2 gases data and reduces the dimension. Subsequently, the model combines VAE and bi-LSTM layers, extracting useful information and predicting the concentration of gases. Traditionally, several statistical models proposed inefficient methods for extracting the data’s implicit feature during the analysis phase (Liu et al. 2004; McKeen et al. 2007). The VAE can extract potential information and reduce data volume, improving the prediction accuracy of the gas concentration. The LSTM is based on a recursive neural network (RNN) to predict the incident over time interval value (Sundermeyer et al. 2012). The bi-LSTM is connected to the hidden layer in both orders between forward and backward. The bi-LSTM receives input from the VAE layer’s output, which generates the essential feature of five gas sensors data. The bidirectional LSTM model is trained using the present and past gas sensor data feature, improving prediction accuracy. Thus, the VAE layer extracts the import feature, and bi-LSTM efficiently predicts the gas concentration in the enclosed region of an underground mine.

Therefore, this study aims to develop a unique model for predicting CH4, CO2, CO, O2, and H2 concentrations, known as the t-SNE_VAE_bi-LSTM model, by learning the in-depth characteristics of these gases. Furthermore, as a result of this study, the proposed model can predict other mine hazards, such as roof fall in underground mines and slope failure in opencast mines with minor modifications.

The following are the paper’s main contributions:

- CH4, CO2, CO, O2, and H2 concentration prediction helps to improve mine safety and early reopening of the sealed-off area to start production.
- The state-of-the-art approach, namely deep learning algorithms, has been used to build a prediction model for CH4, CO2, CO, O2, and H2 gases concentrations.
- An automated metallic gas chamber has been designed that regularly collects gas samples from an underground coal mine sealed-off area at fixed time intervals.
- In the enclosed region of coal mines, a real-time system for monitoring and forecasting CH4, CO2, CO, O2, and H2 gases concentration has been designed.
- The CH4, CO2, CO, O2, and H2 gas concentration prediction model has been proposed by combining t-SNE, VAE, and bi-LSTM models named the t-SNE_VAE_bi-LSTM model.
- The proposed model’s t-SNE method aims to lower the dimension of the collected gas concentration; the proposed model’s VAE layer seeks to retrieve the inner characteristics of low-dimension gas concentration. Finally, the proposed model’s Bi-LSTM layer tries to forecast the concentrations of CH4, CO2, CO, O2, and H2 gases at regular intervals.
- The developed system and prediction model have been validated by deploying the complete unit in a coal mine.

The rest of this paper has been arranged as follows. Section 2 presents a recent survey of the literature on the prediction of gas concentrations. Section 3 provides an overview of the automated and intelligent prediction system. The problem statement for the concentration of gases in the deep coal mine sealed-off region is presented in Sect. 4. Section 5 presents the proposed methodology for predicting real-time gas concentrations. Section 6 presents the detailed dataset, results, and comparative analysis of the different prediction models for gas concentration prediction. Finally, Sect. 7 brings the findings to a conclusion. Figure 1 depicts the paper’s structure.
2 Related works

Mine accidents have increased yearly due to abnormally increased gas concentration levels (Qiao and Zeng 2011; Wang et al. 2013) as underground coal mines extend up to deeper depth. Xia et al. (2016) have mentioned the sudden increase in gas levels and controlling disaster in the coal mine. However, the gas concentration level was changed abnormally before the underground mine accident occurred. Rodriguez et al. (2014) and Song et al. (2019) have also described such a situation and tried to determine the gas concentration of the underground mine but were unable to predict the gas concentration. Different machine learning (ML) methods were also applied by various researchers (Karaca et al. 2006; Xi et al. 2015) to predict gas concentration. There are multiple traditional methods available for predicting the concentration of gases in the underground coal mine. The existing prediction methods can be divided into different categories, like the machine learning forecasting approach (Karaca et al. 2006; Xi et al. 2015), statically prediction approach (Brooks et al. 2016), and mathematical prediction approach (Sundermeyer et al. 2012). Said approaches are not capable of processing vast amounts of data generated from five gas sensors in real time. It is challenging to employ stated techniques in the actual forecasting status of a sealed-off area’s gas level. To predict gas emission of the working face, Ye et al. (2006), Yang et al. (2009), Chen et al. (2016), and Guo et al. (2018) have divided the different gas emission sources. Their plan came from a distinct source and forecasting method. They formed equations for gas emission of other emission areas of a mine, which delivered a significant vision and assistance for gas extraction and ventilation design. However, numerous variables, such as the rate of coal descending per unit of time and the amount of coal that remained in the goaf, were present in this approach (Zheng et al. 2019). Thus, applying the above techniques to forecast future gas levels in the underground coal mine is not easy. To decide the rule of gas concentration trek in the surface of a mine where working is still progress, the researchers observed the equation of gas flow, equation of gas dispersal, and trek equations by a bulky number of mathematical simulations (Cao and Li 2017; Xia et al. 2017). But, mesh separation of the mathematical model was an essential consequence of this experimentation. The simulation result of gas extraction in a mine’s working face depends on a basic model with the perfect margin environment. A mathematical model demonstrated the rule of gas extraction, but it is problematic to accomplish actual forecasting of different gas concentrations.
Many researchers have studied on-time serial sensor data and developed several prediction models. For example, the chaotic time series (CHAOS) model (Zhang et al. 2007; Cheng et al. 2008; Liu 2010), auto-regressive integrated average moving (ARIMA) model (Rokhi et al. 2020), and support vector regression (SVR) model (Kun et al. 2016) proposed for predicting the concentration of gases in underground coal mines. However, as mentioned above, the models’ prediction speed and accuracy were insufficient for actual gas level monitoring. The LSTM neural network is a popular recurrent method for time series data prediction (Lyu et al. 2020). It can remember long-term historical data, which can be used in various applications such as speech recognition, emotion detection, and forecasting detection.

Many researchers have recently developed optimization algorithms. These algorithms described how vital features could be extracted from the gas concentration in underground mines. For example, Abualigah (2019) has proposed a novel classifier for classifying text documents. Further, particle swarm optimization algorithm and dimension reduction technique has been used to get the novel feature in low-dimensional space. This unique characteristic is then employed to boost productivity while lowering the computational cost of the text clustering (TC) technique.

Furthermore, Abualigah et al. (2021a) have proposed a mathematical optimizer to find a solution in an ample search space. It used the distribution behavior of the four basic mathematical operations. Similarly, Abualigah et al. (2021b) have developed a lightweight optimizer motivated by Aquila’s behavior. The optimization procedures of the proposed algorithm are divided into four methodologies for selecting and discovering separate search spaces.

Altabbeeb et al. (2021) have proposed a collaborative modified firefly algorithm to cooperate with the capacitated vehicle routing problem (CVRP). This proposed method efforts to find transport routes with the shortest distance traveled. Abd Elaziz et al. (2021) have proposed modified artificial ecosystem-based optimization (AEO) to solve the task scheduling in the cloud environment. Task scheduling is critical, and optimizing scheduling for IoT task requests can enhance organizational quality and profitability. Hassan et al. (2021) have presented an improved slime mold algorithm (ISMA). They used it to effectively resolve single- and bi-objective financial and emission dispatch (FED) problems while considering valve-point effects. Eid et al. (2021) have developed an improved marine predators algorithm (IMPA) to extend the previous marine predator’s algorithm (MPA). The proposed enhancements result provides faster convergence and avoids possible minima instability for the previous MPA. In addition, IMPA regulates the voltage and current injected into the distributed generation to minimize overall system losses and total voltage deviations. Şahin and Abualigah (2021) have proposed a unique deep neural network-based intrusion prevention technique for determining features using the grouping system. A deep neural network is also utilized to store the time sequence characteristic mapped from actual past data. The proposed approach used an impervious features extraction model to improve the recognition skills of static analyses. Hati et al. (2019) have proposed an intelligent wireless framework to manage the network.

Dey et al. (2021a) and Kumari et al. (2021) have proposed a deep neural network to forecast mining risks and explosive states in the underground mining site. The proposed system indicates the safe condition of the working zone of the underground mines by correlating several hazards parameters. In addition, Dey et al. (2021b) have presented a safe architecture for training the model securely. Similarly, Dey et al. (2021c) have introduced a deep network-based secure communication channel in the mining site to secure communication in underground mines. Muduli et al. (2019) have conducted a comprehensive survey on deploying wireless sensor network technology in the enclosed region of the mine site. It understood the coal mine limits as well as other aspects of the operating zone. Jiang et al. (2018) have used a hazard adjustment approach based on a machine-learning algorithm to anticipate rock bolt incompetence in underground coal mines.

In the predictive training phase, the support vector machine (SVM) algorithm was used for various mine gas concentrations. Zhang et al. (2016) have optimized the weight of the artificial neural of the machine learning model and predicted gas level using the old and disorder principles. Deng et al. (2018) have developed a combined architecture of regression and swarm optimization to estimate the atmospheric pressure of unexpected combustion processes in the goaf region. Qiang and Pu (2018) have presented a technique for predicting short-term electricity supply. The preceding predictions are based on machine learning algorithms and swarm optimization. Zhao et al. (2020) have recommended a wastewater treatment plant based on artificial intelligence (AI). This framework quickly separated the toxic substances from the wastewater. In addition, AI was used to improve the efficiency and data processing in sewerage systems. Osarogiagbon et al. (2020) have presented a trained machine learning approach for the milling process. This approach smoothly recognized the numerous hazardous events that occurred, mainly during the milling process. Finally, Sharafati et al. (2020) have developed an enhanced data mining technique for forecasting effluent sewage’s mean values and predictability.
| Reference | The existing method | Purpose | Pros | Cons |
|-----------|---------------------|---------|------|------|
| Zheng et al. (2019) | It investigated the unpredictable properties of mining gases indices derived from the gas drainage network | Through a chaotic characteristic analysis, it predicts a gas and coal sudden fire calamity cohabitation | It introduces a novel concept and approach for further developing the notion of simultaneous hazard prediction | Prediction accuracy is less compared to other techniques |
| Abualigah (2019) | Krill herd algorithms were proposed for feature selection | The proposed technique is solving the text document clustering problem | It increases textual classification efficiency | Application is limited to feature selection |
| Rekhi et al. (2020) | The forecasting model was proposed using the ARIMA model | It is forecasting the monthly air quality | It provides good accuracy in stationary as well as non-stationary time series data | Less accuracy in long-term modeling |
| Lyu et al. (2020) | LSTM based forecasting model has been adopted with multi-sensor fusion | Short-term prediction of gas concentration | The prediction accuracy is improved using the fusion of the sensor’s data | The complexity of the model is much higher |
| Zhao et al. (2020) | It has proposed a wastewater treatment plant’s process using AI | The model separates the toxic substances in the wastewater | It improves the efficiency and data processing in wastewater treatment plants | Unable to process the data in real-time |
| Osarogiagbon et al. (2020) | Trained a machine learning approach for the milling process | The proposed algorithm identifies the various dangerous happenings, mainly during the drilling process | It is an advancement of learning algorithms for new areas or domains with little data | Unable to process large amount dataset |
| Sharafati et al. (2020) | The optimum wastewater conditions were predicted using a boost regression, gradient boost regression, and random forest regression | The proposed model forecasts the sewage performance characteristics for a commercial wastewater treatment facility | The predictive model achieves acceptable performance | The suggested model is incapable of processing large volumes of data in real-time |
| Abualigah et al. (2021a) | An arithmetic optimization algorithm was proposed for detection | It optimizes operations over a broad range of metaheuristic algorithms | It produces impressive outcomes for addressing complex optimization issues | It solves only the optimization problem |
| Abualigah et al. (2021b) | Proposed an optimizer using aquila optimizer (AO) | It optimizes operations over a broad range of metaheuristic algorithms | It is an efficient metaheuristic algorithm | The algorithm applies only to the optimization problem |
| Atlabeeb et al. (2021) | Proposed a cooperative hybrid firefly algorithm (CHFA) | It searches transport routes with the shortest distance traveled | It prevents the model from falling into the local optima problem | It solves only the optimization problem |
| Abd Elaziz et al. (2021) | Proposed a modified artificial ecosystem-based optimization | It performs task scheduling in the cloud environment | It is a well-known metaheuristic for solving combinatorial optimization problems | The algorithm applies only to the optimization problem |
| Hassan et al. (2021) | Proposed an improved slime mold algorithm (ISMA) | It solves single- and biobjective financial and emission dispatch (FED) problems | The proposed ISMA outperforms SMA, particle swarm optimization (PSO), and tree-seed algorithm (TSA) in solving the FED problem | It solves only the optimization problem |
| Eid et al. (2021) | Developed an improved marine predator’s algorithm (IMPA) to extend the previous MPA | It controls the active and reactive power injected into distribution systems | It enhances the results by convergence and avoids possible marginal instability for the previous MPA | It is employed only for optimization problems |
| Şahin and Abualigah (2021) | Proposed a unique deep network for intrusion recognitions | It is deployed for the automatic detection of software vulnerabilities | It uses a waterproof features extraction model to improve the recognition skills of static analyses | Application is limited to vulnerability detection |
Deep learning has mostly overtaken standard machine learning models in recent times. The deep learning method automatically extracts the vital feature from the sequence of data. Moreover, it compresses the input data using multi-layer applications. Thus, it reduces the over-fitting problem during the model learning time. However, the performance accuracy of the deep neural network is not efficient for complex time series data and is unable to processed massive amounts of data in real time. Hence, the t-SNE-VAE_bi-LSTM model has been proposed in this paper for the accurate and efficient prediction of gases present in underground coal mines in real time. Some of the recent research for predicting hazards and different optimization techniques is summarized in Table 1.

### 3 System description

An automated and intelligent system is designed to monitor the condition of a blaze inside an enclosed region of the underground mining site in real time. The tracking of fire status is achieved by predicting concentrations of gases inside the enclosed region. The sealed-off area is defined as a part of the underground mine where the fire has occurred. The area has been sealed by constructing a wall to control fire by cutting off the oxygen supply and isolate the area from other working faces of the underground mining site. The fire intensity of the isolated area is gradually decreased by cutoff the oxygen flow in this area. Mine fires and explosions take many lives and cause much property
damage every year. However, current methods of analyzing sealed-off areas in mines involve slow, cumbersome manual processes and are prone to error. Thus, an automated gas sampling and predicting system are developed for the sealed-off area to track the status of the fire. The details of the system are depicted in Fig. 2. The system consists of a data acquisition system, WLAN, five gas sensors \((CH_4, CO_2, CO, O_2, \text{ and } H_2)\) fitted inside a box/chamber, two solenoid valves, a suction pump, and a prediction model for collecting gas concentration from the sealed-off area at a fixed time interval. The detailed specification of gas sensors is given in Table 2. The system is connected with a pipe fitting in a fire-stopping brick wall for air sampling purposes from a sealed-off area. After each predefined time interval, the solenoid valve near the pipe opens, and the suction pump starts sucking air from the sealed-off area. The rear solenoid valve also opens subsequently. The sucked air passes through the box/chamber, and gas concentration is measured by different sensors fitted inside the room. It measures the gas concentration and sends it to the prediction model using WLAN. The other gas concentration is predicted using t-SNE and VAE with the bi-LSTM model in the prediction model. After the prediction process, the prediction result is uploaded to the cloud storage using WLAN. The concerned mine management of an underground mine accesses the prediction result from the cloud storage. It continuously monitors the gas level in the enclosed region of the underground mining site. The measurement process continues for 2 min. Then the system stops operation till the next cycle starts.

### 4 Problem scenario

The previously defined models are unable to process the complex and multidimensional time series sensor data effectively. As a result, the models cannot efficiently predict the gas concentration in the enclosed region of the underground mining site. The five fixed time series gas sensors data are collected from the designed metallic gas chamber and uploaded to the prediction model to predict the gas concentration accurately. The accumulated time series data includes 150 days of \(CH_4, CO_2, CO, O_2, \text{ and } H_2\) concentrations, including complex, multidimensional, and noisy data. Hence, there is a requirement for reducing dimension and noise from the collected data. The extracted important feature from collected data has been utilized for efficiently predicting gas concentration. Appropriate deep learning is employed to reduce dimension and noise and extract the vital feature of the multidimensional time series data. Therefore, t-SNE, VAE, and bi-LSTM neural networks have been utilized to predict gas concentration efficiently.

In this case, we explore a 2D matrix \(A(i,j)\) where the \(i\)th row is denoted as a group of different gases, and the \(j\)th column is indicated as a group of timestamps \(T = (t_1, t_2, \ldots, t_d)\) where \(t_1 < t_2 < \cdots < t_d\). For example, the matrix \(a_{i,1}\) is marked as the value of \(CH_4\) concentration at the timestamp \(t_1\).

\[
A(i,j) = \begin{pmatrix} CH_4 & \cdots & H_2 \\ a_{i,1} & \cdots & a_{i,n} \\ \vdots & \ddots & \vdots \\ a_{d,1} & \cdots & a_{d,n} \end{pmatrix}
\]

The first objective is to minimize the dimension \(A(i,j)\) and mapping \(B(i,j)\) using the t-SNE method. The second objective is to extracts the important feature from \(B(i,j)\) and mapping to \(C(i,j)\) using the VAE method. Finally, the bi-LSTM layer predicts the future gas sensor value input as a \(C(i,j)\).

The main problem scenario of the underground coal mines are as follows:

- There is no real-time \(CH_4, CO_2, CO, O_2, \text{ and } H_2\) concentration prediction system available for the enclosed region of the underground mining site to improve mine safety.

#### Table 2 Permissible limit in underground coal mine and specification of different gas sensors

| Gas sensor       | Permissible limit and specification                                                                 |
|------------------|-------------------------------------------------------------------------------------------------------|
| Methane sensor \((CH_4)\) | 1% isolates electricity and 2% remove personnel, 5 to 15% flammability limit (FL) in air, sensor type—irr |
• There is no automated system available to collect gas concentrations from an enclosed region of the underground mining site. Traditionally, the gas concentration was manually collected from the seal-off site using a sampling bag.

• Previously defined models are incapable of processing complex and multidimensional gas concentrations in real-time.

• Earlier described models cannot reduce dimension, and noise as well as cannot extract the vital feature from the multidimensional gas concentration.

Therefore, an automated metallic gas chamber that collects different gases at a fixed time interval has been designed. Section 3 contains a detailed description of the automated system. In addition, a prediction model has been developed based on t-SNE, VAE, and bi-LSTM neural network techniques. The model can process complex and multidimensional gas concentration data in real time and extract vital features that improve mine safety by reducing the noise and dimension of the collected gas concentration.

5 The proposed method for real-time gas concentration prediction

The real-time gas level prediction method is depicted in Fig. 3. It is divided into three parts. The first part of Fig. 3 describes the dimension reduction process of gas sensor data. The second part is the VAE layer, where data are denoised and extracted from the critical feature. The last part is the bi-LSTM-based prediction model, trained, and validated using past and future features input from the VAE layer.

The VAE is a type of deep learning for de-noising the sensor data and extracting the import feature. Various researchers have already applied the VAE for video anomaly detection (Fan et al. 2020), pattern recognition (Ma et al. 2019), and feature learning (Zhang et al. 2019), and it produced a good result. Hence, the present study has employed the VAE layer to extract the potential feature from multiple sensor data. The bi-LSTM is a deep learning model where data processing, classification, and prediction process are performed. The historical and forecasting sensor values are playing an essential role in efficiently predicting gas concentration. The Bi-LSTM model knows short and long-term dependencies without holding duplicate data from both historical and forecast values. Many researchers have already applied bi-LSTM for speech identification (Ogawa and Hori 2017), classification (Zhao et al. 2018), biomedicine (Tutubalina et al. 2018), and sentimental analysis (Chen et al. 2017), which efficiently generated prediction results from time series data. Therefore, the bi-LSTM method is employed for the prediction process. Also, t-SNE, VAE, and bi-LSTM techniques enhanced the prediction accuracy and decreased model training time.

![Fig. 3 Real-time gas concentration prediction model](image-url)
5.1 Preprocessing of input data

Due to the complex, multidimensional, and noisy data sample, it is difficult to directly train the model in the prediction process, generating inaccurate prediction results. As a result, the data preprocessing approach is critical in decreasing data imbalance and improving prediction outcomes.

In this paper, the t-SNE method is adopted nonlinearly by reducing the dimension of gas sensor data. It shrinks time series data by translating the multidimensional spatially neighborhood’s Gaussian distribution to the low-dimensional space. As a result, the t-SNE technique can successfully capture a considerable fraction of local and global structures on a wide scale (Maaten and Hinton 2008). Furthermore, the similarity between multidimensional sensor data points and low-dimensional space is maintained by measuring the Gaussian joint probabilities between two neighboring data points (Fooladgar and Duwig 2018).

Here, we consider multidimensional gas concentration as a two-dimensional matrix $A(1,j) = (a_{t_1}, a_{t_2}, \ldots, a_{t_d}) \in \mathbb{R}^2$ where $t_1 < t_2 < \cdots < t_d$. In $A(1,j)$ a matrix, first row represents the CH$_4$ gas concentration, and the jth column is denoted as the timestamp of the CH$_4$ gas concentration. The detailed descriptions of the matrix are given in Eq. 1. Similarly, we have been represented the dimensionality of the remaining gas concentration. The conditional probability $p_{ij}$ between two neighboring data points $a_i$ and $a_j$ in timestamp $t_i, t_j$ is given by:

$$p_{ij} = \frac{\exp(-\|a_i - a_j\|^2/2\sigma^2)}{\sum_{m\neq i,j} \exp(-\|a_i - a_m\|^2/2\sigma^2)}$$  \hspace{1cm} (2)

where $\sigma$ is the Gaussian variance concerning the central data point $a_k$, and $a_m$ is another neighboring data point in the timestamp $t_m$. When $p_{ij} = 0$, the joint probability $P_{ij}$ of the multidimensional space is determined as:

$$P_{ij} = \frac{(p_{ij} + p_{ji})}{2d}$$  \hspace{1cm} (3)

where $d$ is denoted as a set of data points of multidimensional gas concentration with a different timestamp. The low-dimensional gas concentration is represented as $B(i,j) = (b_1, b_2, \cdots, b_n) \in \mathbb{R}^2$ where $z < Z$. Similar to the above, there is a set of Gaussian variance $\sigma_k$ in the conditional probability $q_{ij}$ to $\frac{1}{\sqrt{2\pi}}$. The joint probability $Q_{ij}$ of low-dimensional space is defined as:

$$Q_{ij} = \frac{(1 + \|b_i - b_j\|^2)^{-1}}{\sum_{m\neq o} (1 + \|b_m - b_o\|^2)^{-1}}$$  \hspace{1cm} (4)

where $b_i, b_j$ are another two neighboring data points at the timestamp $t_m, t_n$. The t-SNE algorithm seeks a low-dimensional $B(i,j)$ that minimizes the mismatch between $P$ and $Q$ in order to make the low-dimensional gas concentration have the identical joint probability distribution as the multidimensional gas concentration. The Kullback-Leibler (KL) divergence between multidimensional and low-dimension gas concentration is used to measure the correlation between $P$ and $Q$. The loss functions $L$ between $P$ and $Q$ is calculated as:

$$L(b_1, b_2, \cdots, b_n) = \sum_k KL(P_k||Q_k) = \sum_k \sum_j P_{kl} \log \frac{P_{kl}}{Q_{kl}}$$  \hspace{1cm} (5)

The loss function $L$ is minimized in the weight updating process using a gradient descent algorithm. The t-SNE algorithm’s gradient is defined as:

$$\frac{\partial L}{\partial b_i} = 4 \sum_j (P_{ji} - Q_{ij})(b_i - b_j)(1 + \|b_i - b_j\|^2)^{-1}$$  \hspace{1cm} (6)

The weight updating of Eq. (6) is derived as:

$$b_i^n = b_i^{n-1} + \eta \frac{\partial L}{\partial b_i} + \alpha(n)(b_i^{n-1} - b_i^{n-2})$$  \hspace{1cm} (7)

where $\eta$ is the learning rate, $\alpha(n)$ is the momentum at iteration $n$. The dimensionally is reduced in each iterated process described in Eq. (7). The dimension reduction process for a multi-dimension time series dataset using the t-SNE method is given in Algorithm 1.
Algorithm 1 The dimension reduction process for multi-dimension time series dataset using t-SNE method

| Input: | Multidimensional gas concentration $A(i,j) \in \mathbb{R}^z$ learning rate ($\eta$), number of iteration ($n$), momentum ($\alpha(n)$). |
|---|---|
| Output: | Low dimension gas concentration $B(i,j) \in \mathbb{R}^z$. |

Start

For $i \leftarrow n$ to 1 do

For $j \leftarrow t$ to $i$ do

Step1 Calculate multidimensional conditional probability $p$ using Equation (2).

Step2 Construct a multidimensional joint probability of matrix $P$ using Equation (3).

Step3 Calculate low dimension joint probability $Q$ using Equation (4) by minimizing the mismatch between $P$ and $Q$.

Step4 Calculates the loss function by using Equation (5) to find the similarity of $P$ and $Q$.

Step5 Iterate Equation (7) until weight is updated for Equation (6).

End For

End

Finally, a low-dimensional gas concentration matrix $B(i,j) \in \mathbb{R}^z$ where $z < Z$ is generated.

5.2 Essential feature extraction using VAE layer

The VAE layer is taking input from preprocessing layer. In this layer, essential features are extracted from the low dimension gas concentration $B(i,j) \in \mathbb{R}^z$. The extracted dataset is passed to the bi-LSTM layer, which predicts gas concentration. The VAE is a generic deep learning model. The working principle of VAE is similar to variational Bayesian learning. The VAE extracts essential features from a low-dimensional gas concentration and produces new information. Figure 4 represents the VAE layer. According to Fig. 4, the VAE layer is split into two components: encoder and decoder. The encoder is created the latent vector from the input dataset, which extracted the main feature. The decoder rebuilds the input dataset using the latent vector to back to the original input dataset. The input dataset is denoted as $B(i,j)$; latent vector pointed as $C(i,j)$ which extracted the main features, encoder parameter is represented as $\phi$, and decoder parameter is marked as $\theta$. The encoder is labeled as $q_\phi(C(i,j)|B(i,j))$, and decoder is described as $p_\theta(B(i,j)|C(i,j))$.

The VAE training procedure is described as follows:

i The encoder $q_\phi(C(i,j)|B(i,j))$ takes the input from the input dataset $B(i,j)$. Then encoder is generated a latent vector $C(i,j)$ using two vectors; the means vector $\mu_\phi(B(i,j))$ and the variance vector $\sigma^2_\phi(B(i,j))$.

ii The latent vector $C(i,j)$ is sampled based on Gaussian distribution using the mean vector $\mu_\phi(B(i,j))$ and the variance vector $\sigma^2_\phi(B(i,j))$. The reparameterization trick (Kingma and Welling 2013; Kingma et al. 2015) is used in the sample $C(i,j)$.

iii The decoder $p_\theta(B(i,j)|C(i,j))$ has been reconstructed $B(i,j)$ from the latent vector $C(i,j)$. The decoder’s posterior distribution is assumed to be Gaussian in this case. The decoder may still immediately measure the mean vector $\mu_\theta(C(i,j))$ and the variance vector $\sigma^2_\theta(C(i,j))$ to regenerate the $B(i,j)$.

iv We are using the lower bound of the periphery likelihood $p_\theta(B(i,j))$ for calculating the gradient. Then, the parameter is updated in the backpropagation process.

In this paper, the VAE layer utilizes the Gaussian distribution to generate an essential feature from the input dataset. The encoder ($\phi$) and decoder ($\theta$) constraints are trained by maximizing the periphery likelihood log $p_\theta(B(i,j))$. The log $p_\theta(B(i,j))$ is calculated in the following equation:
The architecture of variational auto-encoder layer

\[
\log p_\theta(B(i,j)) = \log \int p_\theta(B(i,j)|C(i,j))p(C(i,j))dC = \log \int q_\phi(C(i,j)|B(i,j)) \frac{p_\theta(B(i,j)|C(i,j))p(C(i,j))}{q_\phi(C(i,j)|B(i,j))}dC
\]

Finally, the VAE error rate is described as follows:

\[
\zeta(\theta, \phi, B(i,j)) = -\log p_\theta(B(i,j))
\]  

5.3 Prediction layer based on bi-LSTM

In this layer, the real-time prediction result is generated using bi-LSTM. The working principle of the LSTM layer is similar to the recurrent neural network (RNN) model. The LSTM model maintains one hidden layer, followed by a regular feed-forward output layer. The traditional RNN cannot resolve the vanishing gradient and long-standing dependents problem. But LSTM efficiently solves the vanishing gradient and long-standing dependents problem. The long-standing dependents problem is defined as when the time interval is increased for time series data. The learned information cannot connect to significantly past information, leading to the vanishing gradients problem. The historical and future characteristics of time series data are helpful in the prediction process. If the model is built using historical and future time series data characteristics, it efficiently predicts the future concentration of gases. But the hidden layer of LSTM only contains the feature from the historical data. As a result, a bidirectional LSTM model is used in this research. The model is trained using history and future features from the time series data, efficiently predicting the gas concentration. The bi-LSTM prediction model is represented in Fig. 5. The left side of Fig. 5 explains the bi-LSTM architecture, and the right side designates LSTM neural network.

The LSTM comprises an input gate, a forget gate, and an output gate. The logistic nonlinearity \(\sigma\) is included in the
three defined gates. The input gate controls the input data, defining how long data are read from input time series data. The VAE model extracted the feature $C_l(t, t/l)$ of $l$ hours before the timestamp $T$ and passed it to the bi-LSTM model as an input. This paper aims to predict the gas concentration of an enclosed region $N$ hours after the timestamp $T$. Both $l$ and $N$ are preset time intervals. The intake is calculated using the subsequent Equation:

$$i_t = \sigma(U_i h_{t-1} + W_i x_t + b_i)$$ (14)

$$c_t = f_t * c_{t-1} + i_t * \tanh(U_c h_{t-1} + W_c x_t + b_c)$$ (15)

where $\sigma$ is denoted as sigmoid function, $i$, $c$ and $f$ are represented as input gate, cell state vector, and forget gate. The $h_t$ is marked as a hidden state vector of the bi-LSTM neural network. The $U_i$ and $U_c$ are defined as the weighted value of the hidden state of the bi-LSTM neural network. The $W_i$ and $W_c$ are the weighted value of the input gate and
cell state for the input $x_t$ of the bi-LSTM neural network. The $b_c$ is defined as a bias vector and, $c_t$ is defined as a cell state vector.

The middle gate (forget gate) determines how far to overlook the present state data. The forget gate manages some data features from the input data feature. The forget gate is defined as:

$$f_t = \sigma(U_fh_{t-1} + W_fh_t + b_f)$$  \hspace{1cm} (16)$$

where $\sigma$ is denoted as a sigmoid function, the $h_{t-1}$ is indicated the hidden state vector, and the subscript $t$ is represented as the timestamp. The $U_f$ is represented as weighted matrices of the hidden state $h_{t-1}$. The $W_f$ is the weighted value of the forget gate for the input $x_t$ of the bi-LSTM neural network. The $b_f$ is denoted as a bias vector.

The output gate finally predicts gas concentration. The hidden state $h_t$ is represented in the next movement. The output gate is defined as:

$$o_t = \sigma(U_oh_{t-1} + W_ox_t + b_o)$$
$$h_t = o_t \cdot \tanh(c_t)$$  \hspace{1cm} (17)$$

where $\sigma$ is denoted as sigmoid function. The $U_o$ is represented as weighed matrices of the hidden state $h_t$. The $W_o$ is the weighted value of the output gate for the input $x_t$ of the bi-LSTM neural network. The $b_o$ is denoted as a bias vector.

The proposed bi-LSTM neural network model efficiently analyses time interval values than LSTM neural network model. It is an analysis of data in both backward and forward movements. The historical and future time interval value can affect the forecasting of the present value. The feature using past and present time series data can more accurately predict gas concentration. The bidirectional LSTM deep learning model’s parameters can be used in the forecasting process. The information is stored in the backward direction vector. From backward to forward, the bi-LSTM model is improved. Therefore, the

---

**Table 3** Input variables of the dataset

| Input variable        | Unit       |
|-----------------------|------------|
| Methane (CH$_4$)      | Percentage (%) |
| Carbon dioxide (CO$_2$) | Percentage (%) |
| Carbon monoxide (CO) | ppm        |
| Oxygen (O$_2$)        | Percentage (%) |
| Hydrogen (H$_2$)      | ppm        |

**Table 4** Details of the dissemination process of the concentration of gases

| Dataset             | Percentage (%) | Time (h) |
|---------------------|----------------|----------|
| Training dataset    | 80             | 2957     |
| Testing dataset     | 20             | 779      |

Fig. 7 Hourly time interval concentration of CH$_4$, CO$_2$, CO$_2$, and H$_2$ gases (from left to right, top to bottom) from September 12, 2019 to February 12, 2020
combination of backward and forward information enhances the prediction result. Figure 6 shows the training process of the bi-LSTM model. Here time series of different gas sensor data are selected for the model training process. The forward LSTM takes the input from \( t = 1 \) to \( 27 \), and the backward LSTM takes the information from \( t = 27 \) to 1. It predicts gas concentration. The combination of backward and forward LSTM produces efficient and error-free prediction results.

6 Experimental results

6.1 Dataset

The experiment was conducted for an Indian underground coal mining site to determine the prediction accuracy of \( \text{CH}_4 \), \( \text{CO}_2 \), \( \text{CO} \), \( \text{O}_2 \), and \( \text{H}_2 \) gases using the proposed model. The dataset contained the hourly time interval gas concentration of a sealed-off area from September 12, 2019, to February 12, 2020, depicted in Fig. 7. The X-axis of Fig. 7 represents one hour’s time interval from September 12, 2019 to February 12, 2020, and Y-axis represents the corresponding concentration values of different gases. Table 3 displays the names and units of the input variables.

The t-SNE_VAE_bi-LSTM, ARIMA, and CHAOS models were trained using the collected data. After model training, different gas concentrations were predicted using the proposed trained model and compared the proposed model’s effectiveness with the existing ARIMA and CHAOS models.

In the experiment process, 80% of the collected data were used for model training, and remaining 20% of data was used to test the system. Table 4 gives the information about the distribution process of the collected dataset. The training dataset of gas concentration was used to train the model. The model was trained using 300 iterations. In each iteration process, the performance of the model was optimized by updating the gradient error. In the model’s validation process, the training parameters were adjusted to increase the trained model’s generalization capability and remove the over-fitting problem by dropping some training parameters. Figure 8 represents the training loss and validation loss of t-SNE_VAE_bi-LSTM of the trained model. Finally, the testing process demonstrated the effectiveness of the trained model, which is described in the result section.

6.2 Prediction result

The experiment was conducted on the proposed prediction model for evaluating the prediction accuracy. The proposed

![Figure 8](image)

**Figure 8** The training and validation losses of the proposed model

| Table 5 | Training parameters used in VAE and bi-LSTM models |
|---------|-----------------------------------------------------|
| Model   | Training parameter                                |
|         | Batch size  | Epochs | Hidden layer | Learning rate | Dropout | Optimizer | Activation function |
| VAE     | 64          | 300    | 10           | \( 10^{-3} \) | 0.5     | Adam      | ReLU                |
| bi-LSTM | 64          | 300    | 50           | \( 10^{-3} \) | 0.5     | Adam      | ReLU                |
(a) Actual concentration versus the forecasted concentration of CH₄

(b) Actual concentration versus the forecasted concentration of CO₂

(c) Actual concentration versus the forecasted concentration of CO
model’s prediction accuracy was compared with two traditional machine learning models, namely ARIMA and CHAOS. Here input gas concentration was passed to the proposed model, and it produced the forecasting gas concentration. The mean squared error (MSE) and mean absolute error (MAE) was used to associate the t-SNE_VAE_bi-LSIM model and two traditional machine learning models, namely ARIMA and CHAOS models. The RMSE and MAE are defined as:

\[
\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{l} (A_i - P_i)^2}{l}} \\
\text{MAE} = \sqrt{\frac{\sum_{i=1}^{l} |A_i - P_i|}{l}}
\]

where \(A_i\) is the actual value of \(\text{CH}_4, \text{CO}_2, \text{CO}, \text{O}_2, \text{and H}_2\) gas concentration, \(P_i\) is denoted as predicted results of five gases and \(i\) is indicated the number of gases. Therefore, the smaller RMSE and MAE values represent the forecasting accuracy and efficiency of the prediction model.

Before the model training process, each gas concentration was preprocessed, as described in Sect. 5.1. After preprocessing, each value of the \(\text{CH}_4, \text{CO}_2, \text{CO}, \text{O}_2, \text{and H}_2\) was normalized to \([0, 1]\). The normalization process is described below:
\[ \text{Normalize}_{\text{Variable}} = \frac{V - \min(V)}{\text{Max}(V) - \min(V)} \]  \hspace{1cm} (21) 

where \( V \) is denoted as gas concentration. After the normalization process, the normalized gas concentrations were sent to VAE with LSTM models for training purposes. Three hundred iterations performed the training process. A backpropagation method was employed to minimize the training error, and the LSTM model was optimized using the Adam optimizer. The batch size was set to 64 samples each iteration throughout the training phase, and the training rate was set at \( 10^{-3} \). Table 5 gives the training parameters used in VAE and bi-LSTM models.

The correlations among CH\textsubscript{4}, CO\textsubscript{2}, CO, O\textsubscript{2}, and H\textsubscript{2} concentrations of gases were verified in the validation process. Then, the prediction model was trained using alone CH\textsubscript{4}, CO\textsubscript{2}, CO, O\textsubscript{2}, and H\textsubscript{2} gas concentration (before correlation) and after correlation of the respective gas concentration. After the training process, the correlated prediction accuracy of the five gases was compared with the proposed t-SNE_VAE_bi-LSTM method with the existing ARIMA and CHAOS machine learning models.

| Date (average of 24 h) | Actual CH\textsubscript{4} value | Proposed model | ARIMA model | CHAOS model |
|-----------------------|----------------------------------|----------------|-------------|-------------|
|                       | Predicted value | % different | Predicted value | % different | Predicted value | % different |
| 13-01-2020            | 2.67 | 2.94 | -10.03 | 2.61 | 2.20 | 2.57 | 3.97 |
| 14-01-2020            | 2.80 | 3.00 | -7.26 | 2.65 | 5.51 | 2.55 | 9.08 |
| 15-01-2020            | 2.90 | 3.05 | -5.25 | 2.68 | 7.64 | 2.58 | 11.08 |
| 16-01-2020            | 2.77 | 3.01 | -8.57 | 2.71 | 2.20 | 2.61 | 5.81 |
| 17-01-2020            | 2.65 | 2.94 | -10.80 | 2.74 | -3.21 | 2.64 | 0.56 |
| 18-01-2020            | 2.77 | 2.99 | -8.19 | 2.77 | -0.03 | 2.67 | 3.59 |
| 19-01-2020            | 2.85 | 3.02 | -6.06 | 2.79 | 1.96 | 2.63 | 7.62 |
| 20-01-2020            | 3.00 | 3.10 | -3.34 | 2.82 | 6.02 | 2.65 | 11.66 |
| 21-01-2020            | 3.00 | 3.11 | -3.57 | 2.84 | 5.21 | 2.69 | 10.21 |
| 22-01-2020            | 3.13 | 3.16 | -0.83 | 2.87 | 8.52 | 2.73 | 12.80 |
| 23-01-2020            | 3.10 | 3.16 | -1.98 | 2.89 | 6.84 | 2.93 | 5.54 |
| 24-01-2020            | 3.14 | 3.17 | -0.90 | 2.91 | 7.42 | 2.95 | 6.14 |
| 25-01-2020            | 3.25 | 3.23 | 0.77 | 2.93 | 10.02 | 2.97 | 8.79 |
| 26-01-2020            | 3.10 | 3.16 | -2.01 | 2.95 | 4.94 | 2.99 | 3.65 |
| 27-01-2020            | 3.26 | 3.23 | 1.01 | 2.96 | 9.02 | 3.00 | 7.79 |
| 28-01-2020            | 3.18 | 3.19 | -0.25 | 2.98 | 6.35 | 3.02 | 5.09 |
| 29-01-2020            | 3.20 | 3.21 | -0.29 | 3.00 | 6.34 | 3.04 | 5.09 |
| 30-01-2020            | 3.28 | 3.24 | 1.14 | 3.01 | 8.14 | 3.07 | 6.41 |
| 31-01-2020            | 3.13 | 3.17 | -1.32 | 3.03 | 3.28 | 3.09 | 1.36 |
| 01-02-2020            | 3.30 | 3.25 | 1.64 | 3.04 | 7.88 | 3.10 | 6.06 |
| 02-02-2020            | 3.50 | 3.33 | 4.81 | 3.05 | 12.77 | 3.09 | 11.68 |
| 03-02-2020            | 3.73 | 3.45 | 7.60 | 3.07 | 17.89 | 2.96 | 20.72 |
| 04-02-2020            | 3.57 | 3.40 | 4.75 | 3.08 | 13.83 | 2.97 | 16.71 |
| 05-02-2020            | 3.36 | 3.30 | 2.00 | 3.09 | 8.16 | 2.99 | 11.14 |
| 06-02-2020            | 3.07 | 3.16 | -2.91 | 3.10 | -0.90 | 3.00 | 2.28 |
| 07-02-2020            | 2.94 | 3.09 | -4.87 | 3.11 | -5.67 | 3.01 | -2.43 |
| 08-02-2020            | 2.77 | 3.01 | -8.68 | 3.12 | -12.69 | 3.03 | -9.34 |
| 09-02-2020            | 2.57 | 2.91 | -13.39 | 3.13 | -21.83 | 3.04 | -18.29 |
| 10-02-2020            | 2.64 | 2.92 | -10.39 | 3.14 | -18.69 | 3.10 | -17.19 |
| 11-02-2020            | 2.88 | 3.04 | -5.36 | 3.14 | -9.02 | 3.22 | -11.80 |
| 12-02-2020            | 3.11 | 3.15 | -1.24 | 3.15 | -1.38 | 3.23 | -3.95 |

| Standard deviation | 5.05 | 8.88 | 8.89 |
Figure 9 shows the real value versus correlated prediction value of CH$_4$, CO$_2$, CO, O$_2$, and H$_2$ for t-SNE_VAE_bi-LSTM, ARIMA, and CHAOS models over 739 h (from January 13, 2020, to February 12, 2020) from the testing data in the forecasting phase. The X-axis of Fig. 9 denotes the time interval (in an hour), and Y-axis indicates the predicted level of CH$_4$, CO$_2$, CO, O$_2$, and H$_2$ gases. The upper part of each figure’s color bar represents the observed value using different machine learning models. The experimental results indicated that the t-SNE_VAE_bi-LSTM model achieved better accuracy than ARIMA and CHAOS models in the forecasting phase.

Table 6 indicates the standard deviation of the percentage difference of the predicted CH$_4$ gas concentrations.

Table 7  MSE and MAE values of the proposed model, ARIMA and CHAOS models in the forecasting phase

| Variable name | Mean square error (MSE) | Mean absolute error (MAE) |
|---------------|-------------------------|--------------------------|
|               | t-SNE_VAE_bi-LSTM | ARIMA | CHAOS | t-SNE_VAE_bi-LSTM | ARIMA | CHAOS |
| CH$_4$        | 0.077                  | 0.106 | 0.146 | 0.369                  | 0.489 | 0.510 |
| CO$_2$        | 0.998                  | 1.035 | 1.017 | 1.018                  | 1.082 | 1.091 |
| CO            | 0.077                  | 0.169 | 0.169 | 0.296                  | 0.412 | 0.411 |
| O$_2$         | 0.298                  | 2.179 | 2.190 | 0.58                   | 1.476 | 1.480 |
| H$_2$         | 0.233                  | 1.468 | 1.433 | 0.549                  | 1.211 | 1.191 |

Figure 10  Standard deviation of percentage difference of the measured and predicted values by the proposed model, ARIMA and CHAOS models

Fig. 11 MSE value of the proposed model, ARIMA and CHAOS models in the prediction process after correlation
for the t-SNE VAE bi-LSTM, ARIMA, and CHAOS models from January 13, 2020 to February 12, 2020. It has 739 h of predicted data from three models, with an average of 24 h of data in each row. The standard deviation has been calculated based on the percentage difference between the actual and predicted CH4 gas concentrations for the three models. Therefore, the standard deviation for the three models has been included in the last row of Table 6. Similarly, in the electronic supplementary material file, Tables A1–A4 have the standard deviation of the three models’ CO2, CO, O2, and H2 concentrations of gases.

Figure 10 depicts a comparison of the standard deviation of CH4, CO2, CO, O2, and H2 gas concentrations for the proposed model, ARIMA, and CHAOS models. For CH4 prediction, the standard deviation of the proposed model, ARIMA, and CHAOS models was found to be 5.05%, 8.88%, and 8.89%, respectively. Similarly, CO2 was found to be 4.48%, 5.65%, and 5.99%; CO was found to be 4.28%, 4.30%, and 4.62%; O2 was found to be 10.85%, 43.0%, and 43.51%; and H2 was found to be 6.94%, 8.49%, and 8.25%. Thus, when compared to the ARIMA and CHAOS models, the proposed model has a lower standard deviation. Consequently, to achieve accuracy, the proposed model beats the ARIMA and CHAOS models.

Mean square error (MSE) and mean absolute error (MAE) results are described for the correlated prediction values of the proposed t-SNE VAE bi-LSTM model with ARIMA and CHAOS models. Table 7 gives MSE and MAE results of the t-SNE VAE bi-LSTM model with ARIMA and CHAOS models. The MSE results of CH4, CO2, CO, O2, and H2 were 0.077, 0.998, 0.077, 0.298 and 0.233, respectively, for the proposed t-SNE VAE bi-LSTM model. The MSE results of CH4, CO2, CO, O2, and H2 were 0.146, 1.017, 0.169, 2.190, and 1.433, respectively, for the CHAOS model. The MAE results of CH4, CO2, CO, O2 and H2 were 0.369, 1.018, 0.296, 0.58 and 0.549, respectively, for the proposed t-SNE VAE bi-LSTM model. The MAE results of CH4, CO2, CO, O2, and H2 were 0.489, 1.082, 0.412, 1.476, and 1.211, respectively, for the ARIMA model. The MAE results of CH4, CO2, CO, O2, and H2 were 0.510, 1.091, 0.411, 1.480, and 1.191, respectively, for the CHAOS model. Figures 11 and 12 represent the comparative analysis of MSE and MAE results of the proposed t-SNE VAE bi-LSTM, ARIMA, and CHAOS models. Figure 11 indicates that the MSE result of the proposed t-SNE VAE bi-LSTM model for the CH4 forecasted value is less than 0.029 and 0.069 for the ARIMA and CHAOS models, respectively. Similarly, the proposed t-SNE VAE bi-LSTM model outperformed the ARIMA and CHAOS models by 0.037 and 0.019 for CO2; 0.092, 0.092 for CO; 1.881, 1.892 for O2; 1.235, 1.200 for H2. Figure 12 shows that the MAE result of the proposed t-SNE VAE bi-LSTM model is less than 0.120 and 0.141 for ARIMA and CHAOS models, respectively, CH4 predicted value. Similarly, the proposed t-SNE VAE bi-LSTM model’s MAE result was 0.064 and 0.073 lower than the ARIMA and CHAOS models for CO2; 0.116 and 0.115 for CO; 0.896 and 0.900 for O2; 0.662 and 0.642 for H2. The prediction accuracy is increased if the value of MSE and MAE is decreased. Figures 11 and 12 clearly show that the t-SNE VAE bi-LSTM model has less MSE and MAE value than ARIMA and CHAOS models. Thus, the proposed t-SNE VAE bi-LSTM model has achieved better prediction accuracy than ARIMA and CHAOS models.

The proposed t-SNE VAE bi-LSTM model was trained using CH4, CO2, CO, O2, and H2 gas concentration from
(a) Actual measured values versus the predicted values before and after correlation for CH$_4$

(b) Actual measured values versus the predicted values before and after correlation for CH$_2$

(c) Actual measured values versus the predicted values before and after correlation for CO

Fig. 13 Actual measured values verses before and after the correlated prediction of CH$_4$, CO$_2$, CO, O$_2$, and H$_2$ concentrations (from top to bottom) using t-SNE VAE bi-LSTM models during the forecasting phase
the dataset and the respective correlation value for each concentration of gases. Figure 13 depicts the actual measured value versus before and after the correlated prediction of CH4, CO2, CO, O2, and H2 gas concentrations using the t-SNE_VAE_bi-LSTM, ARIMA, and CHAOS models over 739 h in the forecasting phase, where the X-axis represents the time interval (h), and the Y-axis represents the individual gas level.

Mean square error (MSE) and mean absolute error (MAE) results are described before correlation and after correlation processes of the proposed t-SNE_VAE_bi-LSTM model. Table 8 gives MSE and MAE results before correlation and after correlation processes of forecasting value of 5 gases. The MSE results of CH4, CO2, CO, O2, and H2 were 0.094, 1.073, 0.077, 0.333, and 0.240, respectively, before correlation. After correlation, the MSE results of CH4, CO2, CO, O2, and H2 were 0.077, 0.998, 0.077, 0.298, and 0.233, respectively. The MAE results of CH4, CO2, CO, O2, and H2 were 0.378, 1.053, 0.296, 0.596, and 0.550, respectively, before correlation. After correlation, the MAE results of CH4, CO2, CO, O2, and H2 were 0.369, 1.018, 0.296, 0.581, and 0.549, respectively. Figures 14 and 15 represent the comparative analysis of MSE and MAE results before and after correlation processes for the t-SNE_VAE_bi-LSTM based forecasting model. Figure 14 shows that MSE results after the correlation process were 0.017, 0.075, 0.000, 0.035, and 0.007 less than before the correlation for CH4, CO2, CO, O2, and H2 forecasting results, respectively, by t-SNE_VAE_bi-LSTM model. Figure 15 depicts that MAE results after the correlation process were 0.009, 0.035, 0.000, 0.015, and 0.001, less than before the correlation for CH4, CO2, CO, O2, and H2 forecasting results, respectively, by t-SNE_VAE_bi-LSTM model. The fewer MSE and MAE results increased the efficiency of the t-SNE_VAE_bi-LSTM model in the forecasting phase. Figures 14 and 15 clearly show that MSE and MAE results after correlation were less than before correlation. Thus, after correlating five gas concentrations, the t-SNE_VAE_bi-LSTM model has achieved better accuracy in the prediction process.
7 Conclusions

A novel prediction technique has been proposed to predict the CH₄, CO₂, CO, O₂, and H₂ concentration of gases in the enclosed region in the underground coal-mining site. The said five gas values had been correlated during training phases. The correlation among five gas concentrations increased the forecasting precision of the proposed model. Before the training process, the five gas concentrations were preprocessed using the t-SNE algorithm, reducing the dimension of the gas concentrations. The preprocess input data were given to the VAE layer, where the essential features were extracted from the input data. It has improved the efficiency of the prediction model. The output data of VAE were sent to the bi-LSTM model, where the actual forecasting model was trained to predict the sealed-off area’s gas concentration. The forwarding and backward direction in bi-LSTM efficiently handled the time interval value and increased the forecasting precision. The forecasting value shows that the proposed model has
fewer predicted MSE and MAE values than ARIMA and CHAOS models. Thus, the proposed model may be utilized for online monitoring and predicting concentrations of gases in the enclosed region of the underground coal mining site.

Future works include predicting other mine hazards, like roof fall in underground mines, slope failure in open cast mines, etc.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s00500-021-06261-8.

**Acknowledgements** The authors would like to express their appreciation to Dr. Pradeep K. Singh, Director of the CSIR-Central Institute of Mining and Fuel Research in Dhanbad, India, for publishing this work. The authors would also like to acknowledge the Ministry of Electronics and Information Technology of the Government of India for funding this research under Grant No. 13(8)/2015-CC&BT.

**Data availability** The datasets created and evaluated during the present study have included electronic supplementary material and will be available from the corresponding author upon reasonable request.

**Declarations**

**Conflict of interest** The authors claim that they do not have any conflicts of interest.

**Ethical approval** There are no investigations involving human subjects or animals done by any of the writers in this article.

**References**

Abd Elaziz M, Abualigah L, Attiya I (2021) Advanced optimization technique for scheduling IoT tasks in cloud-fog computing environments. Future Gener Comput Syst 124:142–154

Abualigah LMQ (2019) Feature selection and enhanced krill herd algorithm for text document clustering. Springer, Berlin, pp 1–165

Abualigah L, Diabat A, Mirjalili S, AbdElaziz M, Gandomi AH (2021a) The arithmetic optimization algorithm. Comput Methods Appl Mech Eng 376:113609. https://doi.org/10.1016/j.cma.2020.113609

Abualigah L, Yousri D, AbdElaziz M, Ewees AA, Al-qaness MA, Gandomi AH (2021b) Aquila optimizer: A novel meta-heuristic optimization algorithm. Comput Ind Eng 157:107250. https://doi.org/10.1016/j.cie.2021.107250

Altabeeb AM, Moheen AM, Abualigah L, Ghalib A (2021) Solving capacitated vehicle routing problem using cooperative firefly algorithm. Appl Soft Comput 108:107403. https://doi.org/10.1016/j.asoc.2021.107403

Brooks W, Corsi S, Fienen M, Carvin R (2016) Predicting recreational water quality advisories: A comparison of statistical methods. Environ Model Softw 76:81–94

Cao J, Li WP (2017) Numerical simulation of gas migration into mining-induced fracture network in the goaf. Int J Mining Sci Techno 27(4):681–685

Chaulya SK, Prasad GM (2016) Sensing and monitoring technologies for mines and hazardous areas. Elsevier, USA

Chen L, Wang E, Feng J, Kong X, Li X, Zhang Z (2016) A dynamic gas emission prediction model at the heading face and its engineering application. J Nat Gas Sci Eng 30:228–236

Chen T, Xu R, He Y, Wang X (2017) Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN. Expert Syst Appl 72:221–230

Cheng J, Bai JY, Qian JS et al (2008) Short-Term Forecasting Method of Coalmine Gas Concentration Based on Chaotic Time Series. J China U Min Techno 37(2):231–235

Collobert R, Weston J (2008) A unified architecture for natural language processing: deep neural networks with multitask learning. In: Proceedings of the 25th international conference on machine learning, USA, pp 160–167

Deng J, Lei C, Xiao Y, Cao K, Ma L, Wang W, Laiwang B (2018) Determination and prediction on three zones of coal spontaneous combustion in a gob of fully mechanized caving face. Fuel 211:458–470

Dey P, Chaulya SK, Kumar S (2021a) Hybrid CNN-LSTM and IoT-based coal mine hazards monitoring and prediction system. Process Saf Environ Prot 125:249–263

Dey P, Chaulya SK, Kumar S (2021b) Secure decision tree twin support vector machine training and classification process for encrypted IoT data via blockchain platform. Concurr Comput 33:16. https://doi.org/10.1002/cpe.6264

Dey P et al (2021c) Deep convolutional neural network based secure wireless voice communication for underground mines. J Ambient Intell Hum Comput. https://doi.org/10.1007/s12652-020-02700-w

Eid A, Kamel S, Abualigah L (2021) Marine predators algorithm for optimal allocation of active and reactive power resources in distribution networks. Neural Comput Appl. https://doi.org/10.1007/s00521-021-06078-4

Fan Y, Wen G, Li D, Qiu S, Levine MD, Xiao F (2020) Video anomaly detection and localization via gaussian mixture fully convolutional variational autoencoder. Comput Vis Image Underst 102920

Fooladgar E, Dwig C (2018) A new post-processing technique for analyzing high-dimensional combustion data. Combust Flame 191:226–238

Guo JH, Cheng ZH, Kong WY (2018) Establishment and application of mathematical prediction model of gas emission rate in fully mechanized coal face. Coal Eng 50:109–113

Hassan MH, Kamel S, Abualigah L, Eid A (2021) Development and application of slime mould algorithm for optimal economic emission dispatch. Expert Syst Appl 182:115205. https://doi.org/10.1016/j.eswa.2021.115205

Hati S, Dey P, De D (2019) WLAN based energy efficient smart city design. Microsyst Technol 25(5):1599–1612

He K, et al. (2016) Deep residual learning for image recognition. In: Proceedings of IEEE conference on computer vision and pattern recognition, USA, pp 770–778

Huang CJ, Kuo PH (2018) A deep CNN-LSTM model for particulate matter (PM2.5) forecasting in smart cities. Sensors 18(7):2220. https://doi.org/10.3390/s18072220

Jiang P, Craig P, Crosky A, Maghrebi M, Canbulat I, Saydam S (2018) Risk assessment of failure of rock bolts in underground coal mines using support vector machines. Appl Stoch Models Bus Ind 34(3):293–304

Jo B, Khan RMA (2018) An internet of things system for underground mine air quality pollutant prediction based on azure machine learning. Sensors 18(4):930. https://doi.org/10.3390/s18040930

Karaca F, Nikov A, Alagha O (2006) NN-AirPol: a neural-networks-based method for air pollution evaluation and control. Int J Environ Pollut 28(3–4):310–325

Kingma DP, Welling M (2013) Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114
Kingma DP, Salimans T, Welling M (2015) Variational dropout and the local reparameterization trick. In: Proceedings of advances in neural information processing systems, pp 2575–2583

Kumar A et al (2013) Application of gas monitoring sensors in underground coal mines and hazardous areas. Int J Comput Electr Eng 3(3):9–23

Kumari K et al (2021) UMAP and LSTM based fire status and explosibility prediction for sealed-off area in underground coal mine. Process Saf Environ Prot 146:837–852

Kun L, Ling-Kai Y, Mei-Ling Z, Jian C (2016) Coalmine gas concentration analysis based on support vector machine. In: Proceedings of the IEEE 3rd international conference on information science and control engineering (ICISCE), Beijing, China, pp 257–261

Liu Z (2010) Chaotic time series analysis. Math Probl Eng 1–2:1–31. https://doi.org/10.1155/2010/720190

Liu Y, Park RJ, Jacob DJ, Li Q, Kiluru V, Samat JA (2004) Mapping annual mean ground-level PM2.5 concentrations using multiangle imaging spectroradiometer aerosol optical thickness over the contiguous United States. J Geophys Res Atmos 109(D22):5025. https://doi.org/10.1029/2004JD005025

Lyu P, Chen N, Mao S, Li M (2020) LSTM based encoder-decoder for short-term predictions of gas concentration using multi-sensor fusion. Process Saf Environ Protect 137:93–105

Ma F, Li Y, Zhang C, Gao J, Du N, Fan W (2019) Mcvae: Margin-based conditional variational autoencoder for relation classification and pattern generation. In: Proceedings of the World Wide Web conference, pp 3041–3048

Maaten LVD, Hinton G (2008) Visualizing data using t-SNE. J Mach Learn Res 9:2579–2626

Mandal R et al (2013) Application of programmable logic controller for gases monitoring in underground coal mines. Eng Sci and Techno 3(3):516–522

McKeen S, Chung SH, Wilczak J et al (2007) Evaluation of several Biotrickling filters for biogas sweetening: Oxygen transfer improvement for a reliable operation. Process Saf Environ Prot 92(3):261–268

Saeed N, Alouini MS, Al-Naffouri TY (2019) Towards the internet of underground things: A systematic survey. IEEE Commun Surv Tutor 21(4):3443–3466

Şahin CB, Abualigah L (2021) A novel deep learning-based feature selection model for improving the static analysis of vulnerability detection. Neural Comput Appl. https://doi.org/10.1007/s00521-021-06047-x

Sharafati A, Asadollah SBHS, Hosseinzadeh M (2020) The potential of new ensemble machine learning models for effluent quality parameters prediction and related uncertainty. Process Saf Environ Prot 140:68–78

Song Y, Yang S, Hu X, Song W, Sang N, Cai J, Xu Q (2019) Prediction of gas and coal spontaneous combustion coexisting disaster through the chaotic characteristic analysis of gas indexes in goaf gas extraction. Process Saf Environ Prot 129:8–16

Sundermeyer M, Schlüter R, Ney H (2012) LSTM neural networks for language modeling. In: Proceedings of the 13th annual conference of the international speech communication association, Portland, Oregon September 9–13, USA

Tutubalina E, Miftahutdinov Z, Nikolenko S, Malykh V (2018) Medical concept normalization in social media posts with recurrent neural networks. J Biomed Inf 84:93–102

Wang H, Cheng Y, Yuan L (2013) Gas outburst disasters and the mining technology of key protective seam in coal seam group in the Huainan coalfield. Nat Hazards 67(2):763–782

Xi X, Wei Z, Xiaoguang R, Yijie W, Xinxin B, Wenjun Y, Jin D (2015) A comprehensive evaluation of air pollution prediction improvement by a machine learning method. In: Proceedings of the IEEE international conference on service operations and logistics, and informatics, Tunisia, pp 176–181

Xia T, Zhou F, Wang X, Zhang Y, Li Y, Kang J, Liu J (2016) Controlling factors of symbiotic disaster between coal gas and spontaneous combustion in longwall mining gobs. Fuel 182:886–896

Xia TQ, Zhou FB, Wang XX et al (2017) Safety evaluation of combustion-prone longwall mining gobs induced by gas extraction: A simulation study. Process Saf Environ Prot 109:677–687

Yang YML, Xue YX, Jiao GD et al (2009) Study on pattern of gas emission at fully-mechanized coal face in Liliu mining area. J China Coal Soc 34:1349–1353

Ye Q, Lin BQ, Jiang WZ (2006) The study of methane law in coal mining face. China Min Mag 5:38–41

Zhang JY, Cheng J, Hoo YH et al (2007) Forecasting coalmine gas concentration based on adaptive neuro-fuzzy inference system. J China U Min Techno 4:494–498

Zhang S, Wang B, Li X, Chen H (2016) Research and application of improved gas concentration prediction model based on grey theory and BP neural network in digital mine. Procedia CIRP 56:471–475

Zhang Z, Jiang T, Zhan C, Yang Y (2019) Gaussian feature learning based on variational autoencoder for improving nonlinear process monitoring. J Proc Control 75:136–155

Zhao Y, Yang R, Chevalier G, Shah RC, Romijnders R (2018) Applying deep bidirectional LSTM and mixture density network for basketball trajectory prediction. Optik 158:266–272

Zhao L, Dai T, Qiao Z, Sun P, Hao J, Yang Y (2020) Application of artificial intelligence to wastewater treatment: a bibliometric analysis and systematic review of technology, economy, management, and wastewater reuse. Process Saf Environ Prot 133:169–182
Zheng CS, Jiang BY, Xue S, Chen ZW, Li H (2019) Coalbed methane emissions and drainage methods in underground mining for mining safety and environmental benefits: a review. Process Saf Environ Prot 127:103–124

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.