Ten Machine Learning Algorithms for Short-Term Forecasting: A Comparative Study in Gas Warning Systems

Robert M.X. Wu[1*#], Niusha Shafiabady[2#], Huan Zhang[3], Haiyan (Helen) Lu[1], Ergun Gide[4], Jinrong Liu[1] & Clement Franck Benoit Charbonnier[1]

[1]Faculty of Engineering and Information Technology, University of Technology Sydney, Australia
[2]Faculty of Science and Technology (Sydney Campus), Charles Darwin University, Australia
[3] School of Geography, Shanxi Normal University, Taiyuan, China
[4]School of Engineering and Technology, Central Queensland University, Sydney, Australia

[*] Corresponding authors Email: mingxuan.wu@uts.edu.au & robert_wumx@hotmail.com
[#] The first two authors contributed the first authorship equally to this paper.

Abstract

This research aims to explore more efficient machine learning (ML) algorithms with better prediction assessments for short-term forecasting. Up-to-date literature shows a lack of research on selecting practical ML algorithms for short-term forecasting in real-time industrial applications. The literature reviews the top-tier publications on ML algorithms used in China's industrial applications and finds 29 algorithms. Among them, ten widely used ML algorithms are tested, including ARIMA, BP_Resilient, BP_SOG, KNN, LP, LSTM, Perceptron, RF, RNN, and SVM. A case study is conducted to compare the prediction error and predictive performance assessments. The top two ranked algorithms for prediction error assessment are ARIMA and LR. RF, SVM, BP_SOG, KNN, RNN, and BP_Resilient follow. Perceptron is the last ranked algorithm. For predictive performance assessment, ten algorithms are divided into five clusters: the best (KNN) with the shortest computational time (0.41683s), better (RF, LR, and SVM) with between 1s and 2s, good (Perceptron, BP_SOG, and BP_Resilient) with between 2s and 3s, worse (ARIMA and RNN) with more than 3s, and the worst (LSTM) with the longest computational time (145.19s). Based on the outcomes of comparative analysis, ten ML algorithms can be finally classified into four categories: optimal, efficient, suboptimal, and inefficient. LR and RF are optimal algorithms. Efficient algorithms include SVM, KNN, and ARIMA. Suboptimal algorithms include BP_SOG, BP_Resilient, RNN, and Perceptron. LSTM is an inefficient algorithm. This research finds different results from previous studies between LSTM, KNN, and SVM. The research outcomes significantly explore different views on the performance of ARIMA, LR, KNN, RF, RNN, and SVM compared to previous studies. They are valuable for further research. Finally, this research contributes 20 research questions for further investigation.

Keywords: Machine Learning, Machine Learning Algorithms, Short-Term Forecasting, Gas Warning Systems, Case Study

Acknowledgments

The research presented in this paper is funded by the Shanxi Coking Coal Project (201809fx03), and the Shanxi Social Science Federation (SSKLZDKT2019053).

1 Introduction

Machine learning (ML) (including deep learning) approaches have been widely used to explore a vast number of predictor variables in prediction ability (Féret et al. 2019, p.2, p.11; Arango, Aristizábal & Gómez 2021, p.993). Plenty of ML algorithms have been used to analyze and harness the power of an enormous amount of information (Chan et al. 2020, p.375).

However, choosing the appropriate feature selection method for a specific scenario is not trivial (Afrash et al. 2023, p.2). Based on the time scale, forecasting can be classified into four categories: very short-term forecasting (a few seconds to 30 minutes ahead), short-term forecasting (30 minutes to six hours ahead), medium-term forecasting (six hours to one-day ahead), and long-term forecasting (one-day to one-week ahead) (Soman et al. 2010, p.2). Up-to-date literature shows that there is a lack of research on the selection of practical ML algorithms for short-term forecasting in real-time industrial applications,

This research aims to explore a more efficient ML algorithm with better prediction assessments for short-term forecasting, which uses the three-hour dataset to predict up to one hour ahead of the dataset. A case study method is applied to this research, which compares the prediction error assessment and predictive performance assessment of ML algorithm to enhance the understanding of the research outcomes. The following sections include literature, methodology, case study, discussions, conclusion, contribution, different views, limitations, other further research, and implications for ML education.

2 Literature

2.1 Literature Review

From a research perspective, China has become a world leader in AI publications and patents (Li, Tong & Xiao 2021). Reviews of China's research in ML algorithms used in industrial applications will assist researchers and practitioners in understanding the current situation of ML approaches. The literature reviews the top-tier publications on ML algorithms used in China's industrial applications between 2016 and 2020.

29 algorithms are found in 347 industrial applications. They include Back-Propagation (BP) (27.38%, 95 out of 347), Support Vector Machine (SVM)(24.50%, 85 out of 347), Linear Regression (LR) (8.65%, 30 out of 347), Perceptron (5.19%, 18 out of 347), Recurrent Neural Networks (RNN) (4.90%, 17 out of 347), Random Forest (RF) (3.75%, 13 out of 347), Convolutional Neural Networks (CNN) (3.17%, 11 out of 347), K-means (3.17%, 11 out of 347), AdaBoost (2.88%, 10 out of 347), Bayesian Network (2.59%, 9 out of 347), K-Nearest Neighbour (KNN) (2.02%, 7 out of 347), Stepwise Regression (1.44% , 5 out of 347), Naive Bayes (1.44%, 5 out of 347), Self-Organizing Map (SOM) (1.15%, 4 out of 347), Partial Least Squares Regression (PLSR) (1.15%, 4 out of 347), Logistic Regression (1.15%, 4 out of 347), Learning Vector Quantization (LVQ) (0.86%, 3 out of 347), Classification And Regression Tree (CART) (0.86%, 3 out of 347), Hierarchical Clustering(0.58%, 2 out of 347), C4.5 (0.58%, 2 out of 347), Radial Basis Function Networks (RBFN) (0.29%, 1 out of 347), Locally Weighted Learning (LWL) (0.29%, 1 out of 347), Projection pursuit (0.29%, 1 out of 347), Principal Component Regression (PCR) (0.29%, 1 out of 347), Partial least squares discriminant analysis (PLS) (0.29%, 1 out of 347), Linear Discriminant Analysis (LDA) (0.29%, 1 out of 347), Gradient Boosted Regression Trees (GBRT) (0.29%, 1 out of 347), Expectation Maximization (0.29%, 1 out of 347), and Ridge Regression (0.29%, 1 out of 347) (see Appendix 1).

Among the above algorithms, nine have been discussed by more than ten publications, including BP, SVM, LR, Perceptron, RNN, RF, CNN, K-means, and AdaBoost (see Figure 1).

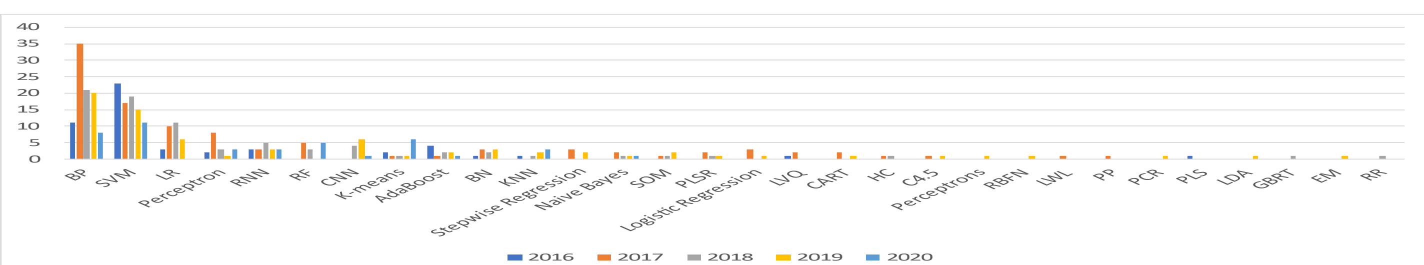


Figure 1: ML Algorithms Used in China's Industrial Applications between 2016 and 2020

2.2 Most-used ML Algorithms

Among the nine most-used ML algorithms, AdaBoost is used for classification and regression tasks (EL Bilali et al. 2021, p.2). The classification method needs a proper training mechanism to be well-applied for prediction tasks. CNN utilizes a convolutional layer to detect patterns in input data for classification or prediction (Readshaw & Giani 2021, p.17354). It is usually used for image processing applications. K-means algorithm works for partitioning the data into the set of clusters defined by centroids and starts with initial estimates for the centroids. These estimates are randomly generated from the datasets (Srikanth, Zahoor Ul Huq & Siva Kumar 2022, p.5). Thus, AdaBoost, CNN, and K-means algorithms are unsuitable for applying to gas warning system applications in this research. They will not be tested in this research.

Besides the above wide-used ML algorithms in China's research, ARIMA is also a popularly used algorithm in international research (Brownlee 2018), which is a common approach used for addressing short-term prediction problems in many studies (Kück & Freitag 2021, p.2; Pu et al. 2021, p.38). ARIMA can account for underlying trends, autocorrelation, and seasonality and allows for flexible modeling of different types of impacts (Schaffer, Dobbins & Pearson 2021, p.11). However, ARIMA cannot effectively capture all the details in very short-term forecasting (Aasim, Singh & Mohapatra 2019, p.766).

BP is one of the most widely used neural networks developed originally for networks of neuron-like units (Rumelhart, Hinton & Williams 1986, p.533). Because of its simple structure, BP can effectively solve the approximation problem of nonlinear objective functions, such as system simulation, function fitting, pattern recognition, and other fields (Huang et al. 2020, p.5645). This research tests BP_Resilient and Second Order Gradient BP (BP_SOG) as training algorithms in this research. The main reason should be that when the large network topology is selected, the standard BP algorithms have problems, such as getting trapped in a local minimum and slow convergence due to the gradients with atomic magnitude (Erkaymaz 2020, p.16279). It is also believed that BP_Resilient has relatively high accuracy, robustness, and convergence speed (Sui Kim et al. 2020, p.15).

Although KNN is only adopted in a few of China's industrial applications (2.02%, 7 out of 347), this research prefers to test its performance. The main reason should be that KNN is simplistic in its workings and calculations (Uddin et al. 2022, p.2). KNN can bypass the complex equation-solving process with computational efficiency (Dong, Ma & Fu 2021, p.4; Kück & Freitag 2021, p.19) and efficiently work on forecasting accuracy in a wider variety of datasets (Kück & Freitag 2021, p.19; Uddin et al. 2022, p.2)- sometimes without any loss of accuracy (Cunningham & Delany 2022, p.22). As a non-parametric and supervised learning classifier, KNN uses proximity to make classifications or predictions about the grouping of an individual data point (Dritsas & Trigka 2022, p.11) and focuses on the correlation by utilizing raw data characteristics (Dong, Ma & Fu 2021, p.4). It has been widely used in forecasting applications of economics, finance, production, and natural systems (Kück & Freitag 2021, pp.2-3).

LR comes under the supervised Learning technique and is one of the most fundamental algorithms in statistics and ML-related fields (Alhakamy et al. 2023, p.2; Dritsas & Trigka 2022, p.10; Mazumder & Wang 2023, p.1226). The mathematical infrastructure of LR is not complex (Şahin et al. 2023, p.4906). Therefore, it is a powerful tool for various tasks in computer vision (Li et al. 2023, p.732) and widely used for predicting and estimating the categorical dependent variable using a given set of independent features (Alhakamy et al. 2023, p.2; Dritsas & Trigka 2022, p.10).

LSTM is also tested in this research. The reason is that although it has not been reported in China's industrial applications until 2020, it is well-known for text classification (Butt et al. 2023, p.3040) and has more frequently been used for forecasting than other algorithms (Elsaraiti & Merabet 2021, p.15). LSTM is a special kind of RNN (Butt et al. 2023, p.3055; Mahmoud et al. 2022, p.405; Van Houdt, Mosquera & Nápoles 2020, p.5931), which may overcome the exploding/vanishing gradient problems that typically arise when learning long-term dependencies, even when the minimal time lags are very long (Sherstinsky 2020, p.12; Van Houdt, Mosquera & Nápoles 2020, p.5931).

Perceptron is one of the most straightforward ANN architectures (Sharma, Kim & Gupta 2022, p.5) and the most typical type of neural predictive network (Moayedi et al. 2021, p.4). Perceptron is designed to approximate any continuous function and can use any arbitrary activation function (Dritsas & Trigka 2022, p.11). It can solve problems that are not linearly separable (Cabeza-Ramírez et al. 2022, p.13; Dritsas & Trigka 2022, p.11) and be applied to produce efficient solutions to problems of overwhelming complexity for conventional computing methods such as holes through edges and through points (Calude, Heidari & Sifakis 2023, p.844; Li et al. 2022, p.2).

RF is a tree-based algorithm based on the creation of several decision trees that belong to the supervised learning technique (Dritsas & Trigka 2022, p.11; Mahmoud et al. 2022, p.404; Pacheco et al. 2021, p.7). RF is used in classification and regression problems (Dritsas & Trigka 2022, p.11) and for addressing short-term prediction problems (Pu et al. 2021, p.38). RF has the robustness to data of any distribution from a large number of features and could ascertain non-linear effects and complex interactions without prior specification (Smithies et al. 2021, p.2, p.8), which may obtain a more accurate and stable forecast (Pacheco et al. 2021, p.7). One limitation when considering the power of RF is that feature measures can show bias when features are correlated (Smithies et al. 2021, p.8).

RNN is one of the most powerful algorithms for processing sequential data such as time series (Elsaraiti & Merabet 2021, p.2; Wang et al. 2022, p.463). It can predict future multiple-time steps (Wang et al. 2022, p.463) and is considered a competing alternative to forecasting time series (Šestanović & Arnerić 2021, p.10).

SVM is one of the most practical parts of statistical learning theories (Huang et al. 2020, p.5465) and is primarily used for classification problems (Dritsas & Trigka 2022, p.10). Recent studies show SVM is capable of producing predictions of high accuracy (Essam et al. 2022, p.3884), which creates the best line or decision boundary that can segregate n-dimensional space into classes so that it can be easily put the new data point in the correct category in the future (Dritsas & Trigka 2022, p.10). However, the SVM's predictive ability is negatively affected when the utilized data set is significantly noisy, as SVMs are sensitive to noise (Essam et al. 2022, p.2).

Thus, this research focuses on ten algorithms, including ARIMA, BP_Resilient, BP_SOG, KNN, LP, LSTM, Perceptron, RF, RNN, and SVM.

3 Methodology

3.1 Research Flowchart

This research uses four-step processes to find an efficient ML algorithm with better prediction assessments for short-term forecasting. They include data collection and data preparation, prediction error assessment, predictive performance assessment, and comparative analysis (see Figure 2).

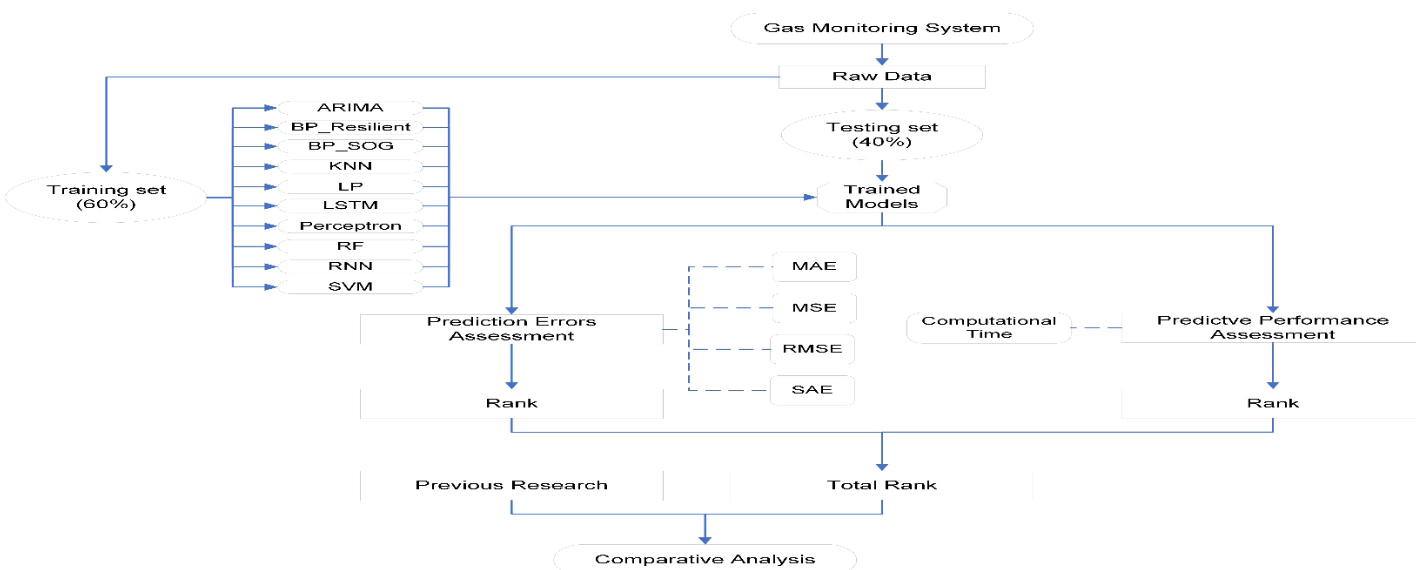


Figure 2: Research Flowchart

3.2 Research Steps

Based on the above Figure 2, four steps for conducting this research can be explained as follows:

- Step 1: Data Collection and Data Pre-processing

Data pre-processing. Data pre-processing is necessary before data analysis since the raw data gathered in most industrial processes usually come with many dataset issues, such as out-of-range values, outliers, missing values, etc. They are not involved in this research. Other data quality issues—such as errors in measurement, noise, missing values, etc.- might be impacted by hardware relocation, sensor removal, added detectors, and/or not in-used sensors, which are not also discussed in this research (Wu et al. 2023, p.3). This research directly obtains raw sensor data into a matrix form fed to the AI prediction system.

Aiming to evaluate these algorithms' performance, each dataset is split into training and testing subsets with a ratio of 60%:40%. The test data are utilized to examine the transferability and predictive capability of the tested algorithms on new data (Mahmoud et al. 2022, p.404). More testing subsets adopted are expected to improve the verification of the test results.

- Step 2: Prediction Errors Assessment

The prediction errors and predictive performance assessment of the employed modeling approaches are subjected to analysis through the use of different statistical indicators to measure the prediction error quality and evaluate the performance of the ML algorithms during the training and testing periods (Ameer et al. 2022, p.8; Sukawutthiva, Sathirapatva & Vongpaisarnsin 2021, p.2; Verhaeghe et al. 2022, p.16). However, it is challenging for most researchers to select specific efficiency criteria among the current performance metrics (Yaseen 2021, p.15). The reason should be that different error metrics have been used to check the effectiveness of the proposed forecasting model (Aasim, Singh & Mohapatra 2019, p.762).

This research reviews Q1 publications related to the predictive performance assessment of ML algorithms between 2020 and 2023. 45 performance criteria are found (see Appendix 2), including absolute average deviation (AAD), average absolute error (AAE), the area under the curve (AUC), commission Error (CE), cross-entropy, coefficient of variance (CoefVar), dice coefficient (DC), developed discrepancy ratio (DDR), Durbin–Watson statistic (DW), error improving rate (EIR), generalization ability (GA), Gain rate criterion (GRC), Gini index (GI), interquartile range and range (IRR), index of agreement (IoA), Kling-Gupta efficiency (KGE), mean absolute error (MAE), mean absolute deviation (MAD), mean absolute percentage error (MAPE), average bias error (MBE), median absolute percentage error (MdAPE), median percentage error (MdPE), mean error (ME), mean square error (MSE), Nash-Sutcliffe efficiency (NSE), omission error (OE), out of bag (OOB) error, overall accuracy (OA), coefficient of determination (R2), relative absolute error (RAE), ranking mean (RM), root mean square of the successive differences (RMS), root mean squared error (RMSE), receiver operating characteristic (ROC) curve, RR variance (RR), sum of absolute errors (SAE), Se/Sy, standard error of prediction (SEP), symmetric MAPE (SMAPE), scatter index (SI), sum of squared errors (SSE), transition matrix features (TMF), t-statistic test (Tstat). It clearly shows that no single or standard error evaluation criteria are adopted as the expected performance method for evaluating the error characteristics of ML algorithms.

Appendix 2 indicates that RMSE (27), MAE (24), R2 (22), and MSE (17) are the most used metrics for evaluating ML algorithms between 2020 and 2023. The results support another recent study by Yaseen (2021, p.15), stating R2, RMSE, and MAE to be the majorly used for the prediction evaluation and Alhakamy et al. (2023, p.8), highlighting MAE, MSE, and RMSE as the primary metrics used to evaluate the performance criteria. However, R2 is oversensitive to extreme values and insensitive to the proportional difference between "actual and predicted values" (Yaseen 2021, p.15). R2 is not, therefore, used in this research.

Among them, MAE is the most straightforward measure to understand and commonly used to interpret linear algorithms (Alhakamy et al. 2023, p.8) and the preferred measure of average model error (Robeson & Willmott 2023, p.2). The smaller the calculated value, the better the performance of the developed algorithm (Mahmoud et al. 2022, p.405). MSE is one of the most prominent criteria in training neural networks and has been employed in numerous learning problems (Heravi & Abed Hodtani 2018, p.6252) and typically interpreted as consistent over- and/or under-prediction of the observations by the model (i.e., the model has non-zero mean bias and/or the regression slope is not one) (Robeson & Willmott 2023, p.2). The value of MSE shows the difference between the predicted and observed values of a model -if there are no errors, it is zero (Alhakamy et al. 2023, p.9). A value of MSE close to zero indicates better performance (Barashid, Munshi & Alhindi 2023, p.8). RMSE is used to measure the concentration of the data in the optimal fit (Alhakamy et al. 2023, p.9) and presents the square root of the variance of the residuals (Cabeza-Ramírez et al. 2022, p.8), As RMSE quantifies the average error and quadratically penalizes significant errors it shows the extent to which the residuals are spread out across the data points in the regression line and the goodness of a model when compared to the actual results and determines the best feature selection algorithms (Alhakamy et al. 2023, p.9, p.12; Verhaeghe et al. 2022, p.16). The value of RMSEs closer to zero indicates better performance (Barashid, Munshi & Alhindi 2023, p.8; Chung et al. 2020, p.8). Although SAE is only used by Yaseen 2021, p.2, p.16), this research believes it is essential to sum up all errors to evaluate the algorithms' quality. SAE is therefore selected for testing algorithms in this research.

Thus, four performance metrics - MAE, MSE, RMSE, and SAE – are used to measure the prediction errors of the employed modeling.

- Step 3: Predictive Performance Assessment

Predictive performance assessment should be another important aspect for evaluating the computational effectiveness of ML algorithms. Computational time is used to measure the predictive performance assessment in this research. The smaller the computation time (the calculated value), the better the performance of the developed algorithm (Mahmoud et al. 2022, p.405).

- Step 4: Comparative Analysis

Based on the above rankings of prediction error assessment on MAE, MSE, RMSE, and SAE and the ranking of computational time, a comparative analysis is conducted compared to the previous research.

4 Case Study

Shanxi Fenxi Mining ZhongXing Coal Industry Co. Ltd (ZhongXing) is selected as a Case Study mine, which is wholly owned by Shanxi Coking Coal Group Co. Ltd - a 485th in the 2020 Fortune Global 500 company located in China (Wu et al. 2021, p.3179). All experiments of ML evaluation are conducted using a standard computer with CPU (11th Gen Intel i7-1165G7 @ 2.80GHZ 2.80GHZ), RAM (16.0 GB), and a 64-bit operating system.

4.1 Data Collection and Data Preparation

The existing gas monitoring system in the Case Study mine comprises two sub-systems: the alarming sub-system used for detecting real-time data obtained from gas sensors and the gas monitoring system used for alarming the safety-responsive team if the gas data outputs reach the threshold limit value (TLV) (Wu et al., 2023a, p.3). The current monitoring sub-system monitors data obtained from gas sensors, temperature sensors, wind sensors, dust sensors, O2 sensors, CO sensors, and CO2 sensors. Data outputs are communicated to a gas monitoring system (Wu et al. 2023a, p.3). The Case Study mine is seeking more responsive ML algorithms for developing a gas early-warning system to predict methane gas (called gas in this paper) concentration (Wu et al. 2022, p.21; Wu et al., 2023, p.1).

This research focuses on gas data. Datasets are collected initially every 15 seconds from a real-time gas monitoring system in the Case Study mine. The gas monitoring system produced four data per minute, 240 per hour, and 5,760 daily. 28,697 valuable datasets were initially acquitted for gas sensor T050401 between 16 April and 16 May 2022.

The datasets are divided into two subsets: 60% for training and 40% for testing datasets. Modeling relations between inputs and outputs are conducted using the above ten algorithms, which use the three-hour dataset to predict up to one hour ahead of the dataset (see Appendix 3).

4.2 Prediction Error Assessment

Four performance metrics (MAE, MSE, RMSE, and SAE) are tested to measure the prediction errors of the employed modeling for both the training and testing datasets (see Table 1).

Performance of each algorithm on criteria with MAE, MSE, RMSE, and SAE to the training dataset shows ARIMA with 0.0043215, 0.018554, 0.13621, and 74.408, BP_Resilient with 0.14471, 0.88631, 0.94144, and 42900000, BP_SOG with 0.071226, 0.81667, 0.9037, and 21100000, KNN with 0.017083, 0.093023, 0.305, and 294.13, LSTM with 0.056083, 0.057971, 0.24077, and 2383.1, LR with 0.0043219, 0.018554, 0.13621, and 74.414, Perceptron with 0.8956, 1.1427, 1.069, and 17218, RF with 0.002815, 0.007159, 0.08461, and 48.464, RNN with 0.067478, 0.83028, 0.9112, and 2533.5, and SVM with 0.004578, 0.018769, 0.137, and 78.823.

Performance of each algorithm on criteria with MAE, MSE, RMSE, and SAE to the testing dataset shows ARIMA with 0.00151, 0.000009, 0.003048, and 17.329, BP_Resilient with 0.060858, 0.006978, 0.083532, and 8020000, BP_SOG with 0.030222, 0.001891, 0.043483, and 3980000, KNN with 0.025468, 0.16305, 0.4038, and 292.35, LSTM with 0.029636, 0.018449, 0.13583, and 1526.4, LR with 0.00151, 0.000009, 0.003048, and 17.333, Perceptron with 0.87563, 0.76756, 0.87611, and 11479, RF with 0.001944, 0.000376, 0.01939, and 22.312, RNN with 0.067697, 0.005384, 0.073375, and 2178.5, and SVM with 0.002069, 0.000011, 0.0032586, and 23.75.

Table 1: Prediction Errors Assessment

Model	Datasets	MAE	MSE	RMSE	SAE
ARIMA	Training	0.004322	0.018554	0.136210	74.408
	Testing	0.001510	0.000009	0.003048	17.329
BP_Resilient	Training	0.144710	0.886310	0.941440	42900000.000
	Testing	0.060858	0.006978	0.083532	8020000.000
BP_SOG	Training	0.071226	0.816670	0.903700	21100000.000
	Testing	0.030222	0.001891	0.043483	3980000.000
KNN	Training	0.017083	0.093023	0.305000	294.130
	Testing	0.025468	0.163050	0.403800	292.350
LR	Training	0.004322	0.018554	0.136210	74.414
	Testing	0.001510	0.000009	0.003048	17.333
LSTM	Training	0.056083	0.057971	0.240770	2383.100
	Testing	0.029636	0.018449	0.135830	1526.400
Perceptron	Training	0.895600	1.142700	1.069000	17218.000
	Testing	0.875630	0.767560	0.876110	11479.000
RF	Training	0.002815	0.007159	0.084610	48.464
	Testing	0.001944	0.000376	0.019390	22.312
RNN	Training	0.067478	0.830280	0.911200	2533.500
	Testing	0.067697	0.005384	0.073375	2178.500
SVM	Training	0.004578	0.018769	0.137000	78.823
	Testing	0.002069	0.000011	0.003259	23.750

Table 2 shows the MAE, MSE, RMSE, and SAE ranks among ten algorithms. For MAE, both ARIMA and LR have the lowest errors (0.00151). Others are followed by RF (0.001944), SVM (0.002069), KNN (0.025468), LSTM (0.029636), BP_SOG (0.030222), BP_Resilient (0.060858), and RNN (0.067697). Perceptron has the worst error (0.87563). For MSE, ARIMA and LR have the lowest errors (0.000009). Others are followed by SVM (0.000011), RF (0.000376), BP_SOG (0.001891), RNN (0.005384), BP_Resilient (0.006978), LSTM (0.018449), and KNN (0.16305). Perceptron has the worst error (0.76756). For RMSE, ARIMA and LR have the lowest errors (0.003048). Others are followed by SVM (0.003259), RF (0.01939), BP_SOG (0.043483), RNN (0.073375), BP_Resilient (0.083532), LSTM (0.13583), and KNN (0.4038). Perceptron has the worst error (0.87611). For SAE, ARIMA has the lowest error (17.329). LR has the second lowest error (17.333). Others are followed by RF (22.312), SVM (23.75), KNN (292.35), LSTM (1526.4), RNN (2178.5), and Perceptron (11479). BP_SOG (3980000) and BP_Resilient have the worst prediction error assessment (8020000).

Based on the MAE, MSE, RMSE, and SAE ranks, the average rank can be seen as ARIMA having a top average rank (1), combining MAE ranked 1, MSE ranked 1, RMSE ranked 1, and SAE ranked 1. LR has the second-top average level (1.3), combining MAE ranked one as the same as ARIMA, MSE ranked one as the same as ARIMA, RMSE ranked one as the same as ARIMA, and SAE ranked 2. RF, SVM, BP_SOG, KNN, LSTM, RNN, and BP_Resilient are followed. Perceptron has the lowest average rank (9.5). Thus, ARIMA and LR are the top two ranked algorithms among all algorithms. RF, SVM, BP_SOG, KNN, RNN, and BP_Resilient are ranked 3 to 9. Perceptron is the last ranked algorithm (10).

Table 2: Rank of Prediction Error Assessment

Model	MAE	Rank	MSE	Rank	RMSE	Rank	SAE	Rank	Average Rank	Rank_Prediction Errors
ARIMA	0.001510	1	0.000009	1	0.003048	1	17.329	1	1.0	1
LR	0.001510	1	0.000009	1	0.003048	1	17.333	2	1.3	2
RF	0.001944	3	0.000376	4	0.019390	4	22.312	3	3.5	3
SVM	0.002069	4	0.000011	3	0.003259	3	23.750	4	3.5	3
BP_SOG	0.030222	7	0.001891	5	0.043483	5	3980000.000	9	6.5	5
KNN	0.025468	5	0.163050	9	0.403800	9	292.350	5	7.0	6
LSTM	0.029636	6	0.018449	8	0.135830	8	1526.400	6	7.0	6
RNN	0.067697	9	0.005384	6	0.073375	6	2178.500	7	7.0	6
BP_Resilient	0.060858	8	0.006978	7	0.083532	7	8020000.000	10	8.0	9
Perceptron	0.875630	10	0.767560	10	0.876110	10	11479.000	8	9.5	10

Based on Table 2, this research finds that MSE and RMSE have the same performance error assessment. The suggestion is that further research does not need to test MSE and RMSE together.

4.3 Predictive Performance Assessment

A predictive performance assessment is followed using computational time testing. The total of training and testing data is used to calculate the time of each ML algorithm. The research outcomes indicate the ten algorithms may be divided into five clusters based on computational time tested – the best (less than 1s), better (between 1s and 2s), good (between 2s and 3s), worse (more than 3s), and the worst (see Table 3).

KNN is the best algorithm with the shortest computational time (0.41683s). RF, LR, and SVM are better algorithms with shorter computational times (less than 2s). RF is the second-best algorithm (1.3503s). LR is the third-best algorithm (1.749s). SVM is also good with a shorter computational time (1.889s). The third cluster includes three algorithms with good computational times to have more than 2s but less than 3s - Perceptron (2.4813s), BP_SOG (2.5108s), and BP_Resilient (2.8363s). Two algorithms are worse with longer computational times (more than 3s) - ARIMA (6.799s) and RNN (34.933s). LSTM is the worst algorithm with the longest computational time (145.19s).

Table 3: Computational Time Testing for Predictive Performance Assessment

Algorithm	Time_ KNN	Time_ RF	Time_ LR	Time_ SVM	Time_ Perceptron	Time_ BP_ SOG	Time_BP _Resilient	Time_ ARIMA	Time_ RNN	Time_ LSTM
Time (s)	0.41683	1.3503	1.749	1.889	2.4813	2.5108	2.8363	6.799	34.933	145.19
Rank_Time	1	2	3	4	5	6	7	8	9	10
Cluster	Best	Better			Good			Worse		Worst

4.4 Comparative Analysis

Based on Table 2 and Table 3 for ranks of prediction errors assessment and computational time, the average rank of all tested algorithms can be thus calculated (see Table 4). The average rank of LR is 2.5 based on the ranks of prediction errors (2) and computational time (3). The average rank of RF is 2.5 based on the ranks of prediction errors (3) and computational time (2). The average rank of SVM is 3.5 based on the ranks of prediction errors (3) and computational time (4). The average rank of KNN is 3.5 based on the ranks of prediction errors (6) and computational time (1). The average rank of ARIMA is 4.5 based on the ranks of prediction errors (1) and computational time (8). The average rank of BP_SOG is 5.5 based on the ranks of prediction errors (5) and computational time (6). The average rank of BP_Resilient is 7 based on the ranks of prediction errors (9) and computational time (5). The average rank of RNN is 7.5 based on the ranks of prediction errors (6) and computational time (9). The average rank of Perceptron is 7.5 based on the ranks of prediction errors (10) and computational time (5). The average rank of LSTM is 8 based on the ranks of prediction errors (6) and computational time (10).

Table 4: Total Rank

Model	Rank_Prediction Errors	Rank_Time	Average Rank	Rank_Total	Category
LR	2	3	2.5	1	**Optimal**
RF	3	2	2.5	1	
SVM	3	4	3.5	3	**Efficient**
KNN	6	1	3.5	3	
ARIMA	1	8	4.5	5	
BP_SOG	5	6	5.5	6	**Suboptimal**
BP_Resilient	9	5	7	7	
RNN	6	9	7.5	8	
Perceptron	10	5	7.5	8	
LSTM	6	10	8	10	**Inefficient**

Thus, ten algorithms can be ranked from 1 to 10 based on the total rank. They can be finally classified into four categories: optimal, efficient, suboptimal, and inefficient algorithms. LR and RF are optimal algorithms with the same average rank (2.5) in the combined assessments of prediction errors and computational time. Efficient algorithms include three algorithms - SVM, KNN, and ARIMA. SVM and KNN have the same rank (3) as efficient algorithms. ARIMA is another efficient algorithm ranked 5. Suboptimal algorithms include BP_SOG ranked 6, BP_Resilient ranked 7; RNN ranked 8, and Perceptron ranked 8. LSTM is an inefficient algorithm, with the worst performance ranked 10.

5 Discussions

This section focuses on each category, including optimal, efficient, suboptimal, and inefficient algorithms, and discusses research findings compared to other studies.

5.1 Optimal Algorithms

Based on Table 4, LR and RF are ranked as optimal algorithms.

- LR

LR is indicated as the optimal algorithm. This research supports other studies on BP_Resilient (Tabbussum & Dar 2020, p.14). LR performs better than BP_Resilient in MSE (0.000009, 0.006978) and RMSE (0.003048, 0.083532). This research also supports other studies on KNN (Pakzad, Roshan & Ghalehnovi 2023, p.8) that LR performs better than KNN in MSE (0.000009, 0.163050) and RMSE (0.003048, 0.403800).

This research has a different view on the performance of LR compared to other studies. For example, recent studies state that LR performs poorly (Patel et al. 2023, p.22) and yields unreliable predictions due to its low flexibility (al-Swaidani et al. 2022, p.3). This research indicates that LR has the second-top average rank (1.3) of combining MAE ranked 1, MSE ranked 1, RMSE ranked 1, and SAE ranked 2. There is a need to verify the prediction error assessment of LR.

This research has different views on prediction error assessment between LR, KNN, and RF compared to other recent studies. A recent study comments that KNN is better than LR in MAE (Pakzad, Roshan & Ghalehnovi 2023, p.8). However, this research indicates that LR (0.001510) is better than KNN (0.025468) in MAE. Two studies affirm that RF has higher discrimination performance and calibrated probabilities than LR (Castonguay et al. 2023, p.45; Mulugeta et al. 2023, p.8). RF is better than LR in MAE, MSE, and RMSE (Pakzad, Roshan & Ghalehnovi 2023, p.8; Šušteršič et al. 2023, p.5). However, a recent study argues LR performs better than RF (Ustebay et al. 2023, p.236). However, this research indicates that LR has a better rank in prediction error assessment than RF in MAE (0.001510, 0.001944), MSE (0.000009, 0.000376), and RMSE (0.003048, 0.019390).

Thus, there is a need to conduct further research and investigate why different results between LR and KNN in MAE and between LR and RF in MAE, MSE, and RMSE are reported by various studies.

- RF

RF is indicated as another optimal algorithm. RF frequently shows statistically lower error performance (Kasbekar et al. 2023, p.10). The finding supports a recent study (Pakzad, Roshan & Ghalehnovi 2023, p.8) that RF is a better assessment than KNN in MAE (0.001944, 0.025468), MSE (0.000376, 0.163050), and RMSE (0.019390, 0.403800). This research also supports two recent studies by Šušteršič et al. (2023, p.5) and Kasbekar et al. (2023, p.10) that RF has only a better prediction of achieving MAE (0.001944) and SAE (22.312) compared to SVM (0.0020690, 23.750).

This research has different views of prediction error assessments between RF and SVM compared to other recent studies. Several recent studies assume that RF may be better than SVM in predicting outcomes in terms of all criteria (Ahmadi, Nopour & Nasiri 2023, pp.18-19; Castonguay et al. 2023, p.45; Hassanzadeh, Farhadian & Rafieemehr 2023, p.13). However, this research indicates that RF and SVM have the same ranks (3) in overall prediction error assessments with MAE (ranked 3 and 4), MSE (ranked 4 and 3), RMSE (ranked 4 and 3), and SAE (ranked 3 and 4) among ten algorithms (see Table 2 and Table 4).

Further research is needed to investigate more measures of prediction error assessments between RF and SVM.

5.2 Efficient Algorithms

Efficient algorithms include three algorithms - SVM, KNN, and ARIMA.

- SVM

This research indicates SVM as one of the most efficient algorithms. This finding supports recent studies by Kasbekar et al. (2023, p.9). Šušteršič et al. (2023, p.5) and Ustebay et al. (2023, p.236). SVM has significantly better prediction achieving MSE (0.000011) and RMSE (0.003259) than RF (0.000376, 0.019390). The finding also supports a recent study by Ustebay et al. (2023, p.236) that SVM significantly outperforms Perceptron in MAE, MSE, RMSE, and SAE, while Perceptron is the last-ranked algorithm in MAE (ranked 10), MSE (ranked 10), RMSE (ranked 10), and SAE (ranked 8) among ten algorithms.

This research finds that previous studies have different views on computational time. Sharma, Kim & Gupta (2022, p.1) affirm that SVM has the shortest training time and prediction speed, while Panesar 2021 (cited in Ustebay et al. 2023, p.236) believes that SVM requires too much training to incur a higher computational cost. This research indicates that SVM has a shorter computational time (1.889s) than Perceptron (2.4813s), BP_SOG (2.5108s), BP_Resilient (2.8363s), ARIMA (6.799s), RNN (34.933s), and LSTM (145.19s). There is a need to investigate further why there is a different research outcome on the computational time of SVM.

- KNN

Although KNN is the best algorithm with the shortest computational time (0.41683s), previous research shows that KNN performs poorly if the training set is large (Cunningham & Delany 2022, p.22; Patel et al. 2023, p.22). This research indicates that KNN has a worse performance of prediction error assessment in MAE (0.025468, ranked 5) and SAE (292.35, ranked 5), while it has the worst MSE (0.16305, ranked 9) and RMSE (0.4038, ranked 9).

This research finds that previous studies have a different view of prediction error assessments between KNN and SVM. The literature shows that KNN outperforms SVM on the majority of the datasets (Mailagaha Kumbure, Luukka & Collan 2020, p.177), while Cunningham & Delany (2022, p.22) believes that KNN may be outperformed by more exotic techniques such as SVM. This research indicates that the performance of MAE, MSE, RMSE, and SAE shows that KNN (0.025468, 0.16305, 0.4038, and 292.35) is worse than SVM (0.002069, 0.000011, 0.003259, and 23.75). Further research is needed to verify the research outcomes on prediction error assessments between KNN and SVM.

This research also has a different view of computational time on KNN compared to other studies. A recent study highlights that KNN algorithms require too much training to incur a higher computational cost (Panesar 2021 cited in Ustebay et al. 2023, p.236). However, this research indicates that KNN has the shortest computational time (0.41683s) compared to other algorithms, from 1.3503s (RF) to 145.19s (LSTM). Further research is needed to investigate the predictive performance assessment of KNN compared to different algorithms.

- ARIMA

This research indicates that ARIMA is another efficient algorithm. However, there is a different view between this research and a previous study on the performance of ARIMA. This research indicates that ARIMA has the highest average rank for prediction error assessment in MAE (0.001510, ranked 1), MSE (0.000009, ranked 1), RMSE (0.003048, ranked 1), and SAE (17.329, ranked 1) among the ten algorithms. However, the literature notes that ARIMA may produce better results with fewer data in previous academic studies' conclusions and worse results with the extensive data in the algorithms generated (Elsaraiti & Merabet 2021, p.15). There is a need to verify the prediction error assessment of ARIMA in further research.

This research has a slightly different view of prediction error assessments between ARIMA and KNN compared to other studies. Previous research states that ARIMA performs marginally better than KNN for the complete set of all-time series (Kück & Freitag 2021, p.19). While this research indicates that ARIMA has the highest performance of prediction error assessment for short-term forecasting rather than KNN, KNN is sixth-ranked for prediction error assessment in MAE (0.025468, ranked 5), MSE (0.163050, ranked 9), RMSE (0.403800, ranked 9), and SAE (292.350, ranked 5). There is a need to investigate the prediction error assessment between ARIMA and KNN in further research.

5.3 Suboptimal Algorithms

Suboptimal algorithms include BP_SOG, BP_Resilients, RNN, and Perceptron. This research does find very few studies that discuss the implementation of BP_SOG, BP_Resilient, and Perceptron compared to other algorithms. There is a need to investigate why they are less applied to China's industrial applications.

This research has a different view of prediction error assessments between RNN and ARIMA compared to previous studies. Previous research believes the superiority of RNN against the traditionally used ARIMA (Šestanović & Arnerić 2021, p.10). However, this research indicates that ARIMA has the highest rank in MAE (ranked 1 with 0.001510), MSE (ranked 1 with 0.000009), RMSE (ranked 1 with 0.003048), and SAE (ranked 1 with 17.329). It is better than RNN in MAE (ranked 9 with 0.067697), MSE (ranked 6 with 0.005384), RMSE (ranked 6 with 0.073375), and SAE (ranked 7 with 2178.500). Therefore, conducting further research to verify the prediction error assessments between RNN and ARIMA is valuable.

5.4 Inefficient Algorithm

This research indicates that LSTM is inefficient with the worst prediction error assessment algorithm and the longest computational time (145.19s).

This research finds that previous studies have different views of the performance of LSTM in the literature. Kasbekar et al. (2023, p.9) states that the statistical comparison results for the absolute errors confirm that LSTM does not perform well on lower errors, which is against other studies, such as LSTM may produce better performances in prediction of modeling time series data (Azeem et al. 2022, p.12; Butt et al. 2023, p.3040; Elsaraiti & Merabet 2021, p.3, p.15; Essam et al. 2022, p.2; Mahmoud et al. 2022, p.405). Thus, it is valuable to investigate further the performance of LSTM.

This research also finds a different view on prediction error assessment between ARIMA and LSTM compared to previous studies. A previous study claims that LSTM outperforms ARIMA with the large quantity of data in

MAE and RMSE criteria (Elsaraiti & Merabet 2021, p.13, p.15). However, this research indicates that ARIMA has the highest rank of overall prediction error assessment in MAE (0.001510) ranked 1, and RMSE (0.003048) ranked 1, while LSTM has MAE (0.029636) ranked 6 and RMSE (0.135830) ranked 8. Thus, further research is needed to verify the prediction error assessment of MAE and RMSE between LSTM and ARIMA.

This research also has a different view on prediction error assessment between LSTM and SVM. A recent study believes that LSTM outperforms SVM (Essam et al. 2022, p.3884). However, this research indicates that SVM has a higher performance with overall rank 3 in prediction error assessment in MAE (0.002069) ranked 4, MSE (0.000011) ranked 3, RMSE (0.003259) ranked 3, and SAE (23.750) ranked 4 rather than LSTM with overall rank 6 in MAE (0.029636) ranked 6, MSE (0.018449) ranked 8, RMSE (0.135830) ranked 8, and SAE (1526.400) ranked 6. Thus, there is a need to conduct further research on prediction error assessment between LSTM and SVM.

6 Conclusion

6.1 Conclusion

This research aims to explore a more efficient ML algorithm with better prediction assessments for short-term forecasting, which uses the three-hour dataset to predict up to one hour ahead of the dataset. Finally, 20 research questions are provided for further investigation.

The literature reviews the top-tier publications on ML algorithms used in China's industrial applications between 2016 and 2020. 29 algorithms are found in 347 applications. Among them, nine algorithms are at least discussed by more than ten publications and applications, including BP (27.38%, 95 out of 347), SVM (24.50%, 85 out of 347), LR (8.65%, 30 out of 347), Perceptron (5.19%, 18 out of 347), RNN (4.90%, 17 out of 347), RF (3.75%, 13 out of 347), CNN (3.17%, 11 out of 347), K-means (3.17%, 11 out of 347), and AdaBoost (2.88%, 10 out of 347) (see Appendix 1). Among the nine most-used ML algorithms, AdaBoost needs a proper training mechanism to be well-applied for prediction tasks. CNN is usually used for image processing applications. K-means algorithm works for partitioning the data into the set of clusters defined by centroids and starts with initial estimates for the centroids. They are not suitable for applying to gas warning systems and will not be tested in this research. ARIMA, BP_Resilient, BP_SOG, and KNN are tested in this research because they are popularly used algorithms for various analyses. Thus, this research focuses on ten algorithms, including ARIMA, BP_Resilient, BP_SOG, KNN, LP, LSTM, Perceptron, RF, RNN, and SVM.

This research reviews Q1 publications related to the predictive performance assessment of ML algorithms between 2020 and 2023. 45 performance criteria are found, including the most-used metrics for evaluating ML algorithms - RMSE, MAE, R2, and MSE (see Appendix 2). However, R2 is oversensitive to extreme values and insensitive and is not used in this research. SAE is therefore selected for testing algorithms in this research because it is essential to sum up all errors to evaluate the algorithms' quality. Thus, four performance metrics - MAE, MSE, RMSE, and SAE – are used to measure the prediction errors of the employed modeling.

A case study method is applied to this research, which compares the prediction assessments of ML algorithm to enhance the understanding of the research outcomes. Datasets are collected initially every 15 seconds from a real-time gas monitoring system in the Case Study mine. The gas monitoring system produced four data per minute, 240 per hour, and 5,760 daily. 28,697 datasets were initially acquitted for gas sensor T050401 between 16 April and 16 May 2022. The datasets are divided into two subsets: 60% for training and 40% for testing datasets. Modeling relations between inputs and outputs are conducted using the above ten algorithms, which use the three-hour dataset to predict up to one hour ahead of the dataset (see Appendix 3).

For prediction error assessment, the research outcomes indicate that ARIMA and LR are the top two ranked algorithms among all algorithms. RF, SVM, BP_SOG, KNN, RNN, and BP_Resilient are ranked 3 to 9. Perceptron is the last ranked algorithm (10) (see Table 2). For predictive performance assessment, the research outcomes indicate the ten algorithms may be divided into five clusters based on computational time tested – the best – KNN - with the shortest computational time (0.41683s), better (RF, LR, and SVM) (between 1s and 2s), good (Perceptron, BP_SOG, and BP_Resilient) (between 2s and 3s), worse (ARIMA and RNN) (more than

3s), and the worst – LSTM - with the longest computational time (145.19s) (see Table 3). Thus, ten ML algorithms can be finally classified into four categories (see Table 4): optimal, efficient, suboptimal, and inefficient algorithms. LR and RF are optimal algorithms. Efficient algorithms include algorithms - SVM, KNN, and ARIMA. Suboptimal algorithms include BP_SOG, BP_Resilient, RNN, and Perceptron. LSTM is an inefficient algorithm with the worst performance.

6.2 Different Views and Further Research

This research finds that MSE and RMSE have the same performance error assessment (see Table 2). The suggestion is that further research does not need to test MSE and RMSE together. This research also finds different results from previous studies on the performance between LSTM, KNN, and SVM. They need to be investigated further as follows:

- Prediction error assessments between KNN and SVM.
- Performance of LSTM.
- Computational time of SVM.

This research significantly explores different views on the performance of ARIMA, LR, KNN, RF, RNN, and SVM compared to previous studies. They are valuable for further research.:

- Computational time on KNN.
- Prediction error assessment of ARIMA.
- Prediction error assessments between ARIMA and KNN.
- Prediction error assessment of LR.
- Prediction error assessment between LR, KNN, and RF.
- Prediction error assessment of MAE and RMSE between LSTM and ARIMA
- Prediction error assessment between LSTM and SVM.
- Prediction error assessments between RF and SVM.
- Prediction error assessments between RNN and ARIMA.

6.2 Limitations and Further Research

The main limitation is that this research aims to find the most suitable ML Algorithms for prediction systems rather than discuss the features of ML Algorithms. Further research needs to investigate the impacts of the advantages and limitations of these algorithms on predicting warning systems. Another limitation is that this research uses data from a Case Study mine in a gas warning system to test ten algorithms to predict gas concentration. There is a need to conduct further research on different industry cases. The third limitation is that this research only focuses on limited prediction error criteria (MAE, MSE, RMSE, and SAE) and predictive performance assessment. It is valuable for testing other prediction error criteria (see Appendix 2).

6.3 Other Further Research

The following research questions also need to be addressed further:

- This research lacks studies on ARIMA BP_Resilient, BP_SOG, and KNN used for China's industrial applications until 2020.
- This research uses the three-hour dataset inputs to predict the following one-hour outputs in a gas warning system in coal mine applications. Testing whether the research outcomes may be verified for very short-term, medium-term, and long-term forecasting is valuable.
- Further research needs to test research outcomes using more prediction error assessments (see Appendix 2) and other performance assessments, such as accuracy.
- Further research must also test whether other industrial applications might verify the research outcomes.
- There is also a need to conduct further research to understand the ML algorithms used for industrial applications between 2021 and 2023.

6.4 Implications for ML Education

The ML algorithms are discussed in this research, and their outcomes should be valuable for IT higher education and professionals to develop up-to-date teaching contexts.

Data availability

The data supporting the study's findings are available in the public domain Figshare with license CC BY4.0 from CC BY4.0 from https://doi.org/10.6084/m9.figshare.24083076.v2.

References

Aasim, Singh, S.N. & Mohapatra, A. 2019, 'Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting', *Renewable energy*, vol. 136, pp. 758-68.

Afrash, M. R., Mirbagheri, E., Mashoufi, M., & Kazemi-Arpanahi, H. 2023, 'Optimizing prognostic factors of five-year survival in gastric cancer patients using feature selection techniques with machine learning algorithms: a comparative study,' *BMC Medical Informatics and Decision Making*, vol. 23, no.1, pp. 54–54. https://doi.org/10.1186/s12911-023-02154-y

Ahmadi, M., Nopour, R. & Nasiri, S. 2023, 'Developing a prediction model for successful aging among the elderly using machine learning algorithms,' *Digital Health*, vol. 9, pp. 20552076231178425. https://doi.org/10.1177/20552076231178425

al-Swaidani A.M., Khwies, W.T., al-Baly, M. & Lala, T. 2022, 'Development of multiple linear regression, artificial neural networks and fuzzy logic models to predict the efficiency factor and durability indicator of nano natural pozzolana as cement additive', *Journal of Building Engineering*, vol. 52, pp. 1-27.

Alhakamy, A., Alhowaity, A., Alatawi, A.A. & Alsaadi, H. 2023, 'Are Used Cars More Sustaina-ble? Price Prediction Based on Linear Regression', *Sustainability (Basel, Switzerland)*, vol. 15, no. 2, pp. 1-17.

Ameer, K., Ameer, S., Kim, Y., Nadeem, M., Park, M., Murtaza, M.A., Khan, M.A., Nasir, M.A., Mueen-Ud-Din, G., Mahmood, S., Kausar, T. & Abubakar, M. 2022, 'A Hybrid RSM-ANN-GA Approach on Optimization of Ultrasound-Assisted Extraction Conditions for Bioactive Component-Rich Stevia rebaudiana (Bertoni) Leaves Extract', *Foods*, vol. 11, no. 6, pp. 1-24.

Arango, M.I., Aristizabal, E. & Gomez, F. 2021, 'Morphometrical analysis of torrential flows-prone catchments in tropical and mountainous terrain of the colombian andes by machine learning techniques', *Natural Hazards (Dordrecht)*, pp. 983-1012.

Azeem, A., Ismail, I., Jameel, S.M., Romlie, F., Danyaro, K.U. & Shukla, S. 2022, 'Deterioration of Electrical Load Forecasting Models in a Smart Grid Environment', *Sensors (Basel)*, vol. 22, no. 12, pp. 1-28.

Barashid, K., Munshi, A. & Alhindi, A. 2023, 'Wind Farm Power Prediction Considering Layout and Wake Effect: Case Study of Saudi Arabia', *Energies (Basel)*, vol. 16, no. 2, pp. 1-22.

Brownlee, J. 2018, 'A Gentle Introduction to Exponential Smoothing for Time Series Forecasting in Python', <https://machinelearningmastery.com/exponential-smoothing-for-time-series-forecasting-in-python/>.

Butt, U.A., Amin, R., Aldabbas, H., Mohan, S., Alouffi, B. & Ahmadian, A. 2023, 'Cloud-based email phishing attack using machine and deep learning algorithm', *Complex & Intelligent Systems*, vol. 9, no. 3, pp. 3043–70. https://doi.org/10.1007/s40747-022-00760-3.

Cabeza-Ramírez, L.J., Rey-Carmona, F.J., Del Carmen Cano-Vicente, M. & Solano-Sánchez, M.Á 2022, 'Analysis of the coexistence of gaming and viewing activities in Twitch users and their relationship with pathological gaming: a multilayer perceptron approach', *Scientific Reports*, vol. 12, no. 1, pp. 1-18.

Calude, C.S., Heidari, S. & Sifakis, J. 2023, 'What perceptron neural networks are (not) good for?', *Information Sciences*, vol. 621, pp. 844-57.

Carvalho, J.E.,S.P., Santos, B.,F., Martins, A.R.,F.A., Braga, S.,L. Pradelle, R.,N.C., Turkovi-cs, F., Perrier, B., Maire, F. & Pradelle, F. 2023, 'Polynomial and ANN models applied to the formation of gums in Brazilian ethanol–gasoline blends—impact of gasoline com-position, ethanol concentration, storage temperature, and aging duration', *Neural Computing & Applications*, vol. 35, no. 22, pp. 16267–84. https://doi.org/10.1007/s00521-023-08396-1

Castonguay, A.C., Zoghi, Z., Zaidat, O.O., Burgess, R.E., Zaidi, S.F., Mueller-Kronast, N., Liebeskind, D.S. & Jumaa, M.A. 2023, 'Predicting Functional Outcome Using 24-Hour Post-Treatment Characteristics: Application of Machine Learning Algorithms in the STRATIS Registry', *Annals of Neurology*, vol. 93, no. 1, pp. 40–9. https://doi.org/10.1002/ana.26528

Chan, S., Reddy, V., Myers, B., Thibodeaux, Q., Brownstone, N. & Liao, W. 2020, 'Machine learning in dermatology: current applications, opportunities, and limitations', *Dermatol Ther*, vol. 10, no. 3, pp. 365–86.

Chung, J., Lee, Y., Jang, W., Lee, S. & Kim, S. 2020, 'Correlation analysis between air temperature and MODIS land surface temperature and prediction of air temperature using tensorflow long short-term memory for the

period of occurrence of cold and heat waves', *Remote sensing (Basel, Switzerland)*, vol. 12, no. 19, pp. 1-25.
Cunningham, P. & Delany, S.J. 2021, 'k-Nearest Neighbour Classifiers - A Tutorial', *ACM computing surveys*, vol. 54, no. 6, pp. 1–25.
Das, P.P., Rabby, M.M., Vadlamudi, V. & Raihan, R. 2022, 'Moisture Content Prediction in Polymer Composites Using Machine Learning Techniques', *Polymers*, vol. 14, no. 20, pp. 1-23.
Debbarma, S. & Ransinchung, G.D. 2020, 'Using artificial neural networks to predict the 28-day compressive strength of roller-compacted concrete pavements containing RAP aggregates', *Road Materials and Pavement Design*, pp. 149-67.
Dong, Y., Ma, X. & Fu, T. 2021, 'Electrical load forecasting: A deep learning approach based on K-nearest neighbors', *Applied Soft Computing Journal*, 99, pp. 1-15.
Dritsas, E. & Trigka, M. 2022, 'Data-Driven Machine-Learning Methods for Diabetes Risk Prediction', *Sensors (Basel, Switzerland)*, vol. 22, no. 14, pp. 1-18.
EL Bilali A., Taleb, A., Bahlaoui, M.A. & Brouziyne, Y. 2021, 'An integrated approach based on Gaussian noises-based data augmentation method and AdaBoost model to predict faecal coliforms in rivers with small dataset', *Journal of hydrology*. vol. 599, pp. 1-11.
Elsaraiti, M. & Merabet, A. 2021, 'A comparative analysis of the arima and lstm predictive models and their effectiveness for predicting wind speed', *Energies*, vol. 14, no. 20, pp. 1-16.
Erkaymaz, O. 2020, 'Resilient back-propagation approach in small-world feed-forward neural network topology based on Newman–Watts algorithm', *Neural computing & applications*, vol. 32, no. 20, pp. 16279–89.
Essam, Y., Huang, Y., Ng, J., Birima, A.H., Ahmed, A.N. & El-Shafie, A. 2022, 'Predicting streamflow in Peninsular Malaysia using support vector machine and deep learning algorithms', *Scientific reports*, vol. 12, no. 1, pp. 1-13.
Fang, W., Zhuo, W., Song, Y., Yan, J., Zhou, T. & Qin, J. 2023, 'Δfree-LSTM: An error distribution free deep learning for short-term traffic flow forecasting', *Neurocomputing (Amsterdam)*, vol. 526, pp. 180-90.
Féret, J.B., le Maire, G., Jay, S., Berveiller, D., Bendoula, R., Hmimina, G., Cheraiet, A., Oliveira, J.C., Ponzoni, F.J., Solanki, T., de Boissieu, F., Chave, J., Nouvellon, Y., Porcar-Castell, A., Proisy, C., Soudani, K., Gastellu-Etchegorry, J. & Lefèvre-Fonollosa, M.J. 2019, 'Estimating leaf mass per area and equivalent water thickness based on leaf optical properties: Potential and limitations of physical modeling and machine learning', *Remote Sensing of Environment*, vol. 231, pp. 110959.
Hammad Khaliq, A., Basharat, M., Talha Riaz, M., Tayyib Riaz, M., Wani, S., Al-Ansari, N., Ba Le, L. & Thi Thuy Linh, N. 2023, 'Spatiotemporal landslide susceptibility mapping using machine learning models: A case study from district Hattian Bala, NW Himalaya, Pakistan', *Ain Shams Engineering Journal*, vol. 14, no. 3.
Hassanzadeh, R., Farhadian, M. & Rafieemehr, H. 2023, 'Hospital mortality prediction in trauma-tic injuries patients: comparing different SMOTE-based machine learning algorithms', *BM-C Medical Research Methodology*, vol. 23, no. 1, pp. 101. https://doi.org/10.1186/s12874-023-01920-w.
Heravi, A. R. & Abed Hodtani, G. 2018, 'A New Correntropy-Based Conjugate Gradient Backpropagation Algorithm for Improving Training in Neural Networks', *IEEE transaction on neural networks and learning systems*, vol. 29, no. 12, pp. 6252–63.
Huang, L., Chen, F., Jhou, M., Kuo, C., Wu, C., Lu, C., Chen, Y., Pei, D., Cheng, Y. & Lu, C. 2022, 'Comparing Multiple Linear Regression and Machine Learning in Predicting Diabetic Urine Albumin–Creatinine Ratio in a 4-Year Follow-Up Study', *Journal of Clinical Medicine*, vol. 11, no. 13, pp. 1-14. https://doi.org/10.3390/jcm11133661
Huang, J., Ko, K., Shu, M. & Hsu, B. 2020, 'Application and comparison of several machine learning algorithms and their integration models in regression problems', *Neural Computing & Applications*, vol. 32, no. 10, pp. 5461–69. https://doi.org/10.1007/s00521-019-04644-5
Kasbekar, R.S., Ji, S., Clancy, E.A. & Goel, A. 2023, 'Optimizing the input feature sets and m-achine learning algorithms for reliable and accurate estimation of continuous, cuffless blood pressure', *Scientific Reports*, vol. 13, no. 1, pp. 7750. https://doi.org/10.1038/s41598-023-34677-9
Kück, M. & Freitag, M. 2021, 'Forecasting of customer demands for production planning by local k-nearest

neighbor models', *International journal of production economics*, vol. 231, pp. 107837.
Li, D.T., Tong, T.W. & Xiao, Y.G. 2021, 'Is China Emerging as the Global Leader in AI? ', *Harvard Business Review*, 18 Feb. https://hbr.org/2021/02/is-china-emerging-as-the-global-leader-in-ai.
Li, F., Fujiwara, K., Okura, F. & Matsushita, Y. 2023, 'Shuffled Linear Regression with Outliers in Both Covariates and Responses', *International Journal of Computer Vision*, vol. 131, no. 3, pp. 732-51.
Li, R. & Song, X. 2023, 'A multi-scale model with feature recognition for the use of energy futures price forecasting', *Expert Systems with Applications*, vol. 211, pp. 118622.
Li, T., Ren, L., Yang, F. & Dang, Z. 2022, 'Analysis of Human Information Recognition Model in Sports Based on Radial Basis Fuzzy Neural Network', *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1-8.
Mahmoud, N., Abdel-Aty, M., Cai, Q. & Yan J.H. 2022, 'Estimating cycle-level real-time traffic movements at signalized intersections', *Journal of intelligent transportation systems*, vol. 26, no. 4, pp. 400–19.
Mailagaha Kumbure, M., Luukka, P. & Collan, M. 2020, 'A new fuzzy k-nearest neighbor classifier based on the Bonferroni mean', *Pattern Recognition Letters*, vol. 140, pp. 172-8.
Mazumder, R. & Wang, H. 2023, 'Linear regression with partially mismatched data: local search with theoretical guarantees', *Mathematical Programming*, vol. 197, no. 2, pp. 1265–303.
Moayedi, H., Gör, M., Kok Foong, L. & Bahiraei, M. 2021, 'Imperialist competitive algorithm hybridized with multilayer perceptron to predict the load-settlement of square footing on layered soils', *Measurement: Journal of the International Measurement Confederation*, vol. 172, pp. 108837.
Moghadam, S.M., Yeung, T. & Choisne, J. 2023, 'A comparison of machine learning models' accuracy in predicting lower-limb joints' kinematics, kinetics, and muscle forces from wearable sensors', *Scientific reports; Sci Rep*, vol. 13, no. 1, pp. 5046.
Mulugeta, G., Zewotir, T., Tegegne, A.S., Juhar, L.H. & Muleta, M. B. 2023, 'Classification of i-mbalanced data using machine learning algorithms to predict the risk of renal graft failures in Et-hiopia', *BMC Medical Informatics and Decision Making*, vol. 23, no. 1, pp. 98. https://doi.org/10.1186/s12911-023-02185-5.
Pacheco, A.d.P., Junior, J.A.d.S., Ruiz-Armenteros, A. & Henriques, R.F.F. 2021, 'Assessment of k-Nearest Neighbor and Random Forest Classifiers for Mapping Forest Fire Areas in Central Portugal Using Landsat-8, Sentinel-2, and Terra Imagery', *Remote sensing (Basel, Switzerland)*, vol. 13, no. 7, pp. 1345.
Pakzad, S.S., Roshan, N. & Ghalehnovi, M. 2023, 'Comparison of various machine learning algorithms used for compressive strength prediction of steel fiber-reinforced concrete', *Scientific Reports*, vol. 13, no. 1, pp. 3646. https://doi.org/10.1038/s41598-023-30606-y.
Panesar, A. 2021, 'Machine learning and AI for healthcare: big data for improved health outcomes (Second edition.) ', *APress*, https://doi.org/10.1007/978-1-4842-6537-6
Patel, S., Wang, M., Guo, J., Smith, G. & Chen, C. 2023, 'A Study of R-R Interval Transition Matrix Features for Machine Learning Algorithms in AFib Detection', *Sensors (Basel, Switzerland)*, vol. 23, no. 7, pp. 3700. https://doi.org/10.3390/s23073700.
Pu, Z., Liu, C., Shi, X., Cui, Z. & Wang, Y. 2021, 'Road surface friction prediction using long short-term memory neural network based on historical data', *Journal of intelligent transportation systems*, vol. 26, no. 1, pp. 34–45.
Rahbar, A., Mirarabi, A., Nakhaei, M., Talkhabi, M. & Jamali, M. 2022, 'A Comparative Analysis of Data-Driven Models (SVR, ANFIS, and ANNs) for Daily Karst Spring Discharge Prediction', *Water resources management*, vol. 36, no. 2, pp. 589–609.
Readshaw, J. & Giani, S. 2021, 'Using company-specific headlines and convolutional neural networks to predict stock fluctuations', *Neural Computing and Applications*, vol. 33, pp. 17353–67.
Robeson, S.M. & Willmott, C.J. 2023, 'Decomposition of the mean absolute error (MAE) into systematic and unsystematic components', *PloS One*, vol. 18, no. 2, pp. 1-8.
Rumelhart, D.E., Hinton, G.E. & Williams, R.J. 1986, 'Learning representations by back-propagating errors', *Nature*, vol. 323, no. 9, pp. 533-6.
Şahin, D.Ö., Kural, O.E., Akleylek, S. & Kılıç, E. 2023, 'A novel permission-based Android malware detection system using feature selection based on linear regression', *Neural Computing & Applications*, vol. 35, no. 7, pp. 4903-18.
Schaffer, A.L., Dobbins, T.A. & Pearson, S.A. 2021, 'Interrupted time series analysis using autoregressive

integrated moving average (ARIMA) models: a guide for evaluating large-scale health interventions', *BMC medical research methodology*, vol. 21, no. 1, pp. 1-12.
Šestanović, T. & Arnerić, J. 2021, 'Can Recurrent Neural Networks Predict Inflation in Euro Zone as Good as Professional Forecasters? ', *Mathematics*, vol. 9, no. 19, pp. 1-14.
Sharma, R., Kim, M. & Gupta, A. 2022, 'Motor imagery classification in brain-machine interface with machine learning algorithms: Classical approach to multi-layer Perceptron model', *Biomedical Signal Processing and Control*, vol. 71, pp. 1-10.
Sherstinsky, A. 2020, 'Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network', *Physica.D: Nonlinear Phenomena*, vol. 404, pp. 132306.
Singh, S.N. & Mohapatra, A. 2019, 'Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting', *Renewable energy*, vol. 136, pp. 758-68.
Smithies, T.D., Campbell, M.J., Ramsbottom, N. & Toth, A.J. 2021, 'A Random Forest approach to identify metrics that best predict match outcome and player ranking in the esport Rocket League', *Scientific reports*, vol. 11, no. 1, pp. 1-12.
Soman, S.S., Zareipour, H., Malik, O. & Mandal, P. 2010, 'A review of wind power and wind sp-eed forecasting methods with different time horizons', , IEEE, pp. 1–8. https://doi.org/10.1109/NAPS.2010.5619586
Sridharan, K. & Shah, S. 2023, 'Developing supervised machine learning algorithms to evaluate t-he therapeutic effect and laboratory-related adverse events of cyclosporine and tacrolimus in ren-al transplants', *International Journal of Clinical Pharmacy*, vol. 45, no. 3, pp. 659–68. https://doi.org/10.1007/s11096-023-01545-5
Srikanth, K., Zahoor Ul Huq, S. & Siva Kumar, A.P. 2022, 'Big data based analytic model to predict and classify breast cancer using improved fractional rough fuzzy K-means clustering and labeled ensemble classifier algorithm', *Concurrency and computation*, vol. 34, no. 10, pp. 1-21.
Sui Kim, I.T., Sethu, V., Arumugasamy, S.K. & Selvarajoo, A. 2020, 'Fenugreek seeds and okra for the treatment of palm oil mill effluent (POME)–Characterization studies and modeling with backpropagation feedforward neural network (BFNN) ', *Journal of Water Process Engineering*, vol. 37, pp. 1-16.
Sukawutthiya, P., Sathirapatya, T. & Vongpaisarnsin, K. 2021, 'A minimal number CpGs of ELOVL2 gene for a chronological age estimation using pyrosequencing', *Forensic science international*, vol. 318, p. 110631.
Šušteršič, T., Gribova, V., Nikolic, M., Lavalle, P., Filipovic, N. & Vrana, N.E. 2023, 'The E-ffect of Machine Learning Algorithms on the Prediction of Layer-by-Layer Coating Properties', *ACS Omega*, vol. 8, no. 5, pp. 4677–86. https://doi.org/10.1021/acsomega.2c06471
Tabbussum, R. & Dar, A.Q. 2020, 'Comparative analysis of neural network training algorithms for the flood forecast modelling of an alluvial Himalayan river', *Journal of Flood Risk Management*, vol. 13, no. 4, pp. 1-18.
Taheri, M., Shamsi Anboohi, M., Nasseri, M. & Mohammadian, A. 2023, 'Developing a dynamic semi-distributed hydrological model considering interactions between soil moisture and evapotranspiration: application of bulk transfer method', *Hydrological Sciences Journal*, vol. 68, no. 2, pp. 228-45
Tataru-Farmus, R., Cimpoesu, R., Nica, I. & Suteu, D. 2023, 'Biosorbent Based on Poly (vinyl alcohol)-Tricarboxi-Cellulose Designed to Retain Organic Dyes from Aqueous Media', *Polymers*, vol. 15, no. 3, pp. 1-16.
Uddin, S., Haque, I., Lu, H., Moni, M.A. & Gide, E. 2022, 'Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction', *Scientific Reports*, vol. 12, no. 1, pp. 6256. https://doi.org/10.1038/s41598-022-10358-x
Ustebay, S., Sarmis, A., Kaya, G. K. & Sujan, M. 2023, 'A comparison of machine learning algorithms in predicting COVID-19 prognostics', *Internal and Emergency Medicine*, vol. 18, no. 1, pp. 229–39. https://doi.org/10.1007/s11739-022-03101-x
Van Houdt, G., Mosquera, C. & Nápoles, G. 2020, 'A review on the long short-term memory model', *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5929-55.
Verhaeghe, J., Dhaese, S.A.M., De Corte, T., Vander Mijnsbrugge, D., Aardema, H., Zijlstra, J.G., Verstraete, A.G., Stove, V., Colin, P., Ongenae, F., De Waele, J.J. & Van Hoecke, S. 2022, 'Development and evaluation of uncertainty quantifying machine learning models to predict piperacillin plasma concentrations in

critically ill patients', *BMC Medical Informatics and Decision Making*, vol. 22, no. 1, pp. 1-17.
Wang, B., Vu, H.L., Kim, I. & Cai C. 2022, 'Short-term traffic flow prediction in bike-sharing networks', *Journal of intelligent transportation systems*, vol. 26, no. 4, pp. 461–75.
Wu, R.M.X., Zhang, Z.W., Zhang, H., Wang, Y.W., Shafiabady, N., Yan, W.J., Gou, J.W., Gide, E. & Zhang, S.Q. 2023, 'An FSV Analysis Approach to verify the Robustness of The Triple-Correlation Analysis Theoretical', *Scientific Reports*, vol, 13, no. 1, pp. 1-20.
Wu, R.M.X., Yan, W.J., Shafiabady, N., Zhang, H., Guo, J.W., Shi, Y., Liu, B., Gide, E., Kang, C.L., Zhang, Z.W., Shen, B., Li, X.Q., Fan, J.F., He, X.Q., Soar, J., Zhao, H.J., Sun, Li., Huo, W.Y. & Wang, Y. 2023a, 'Using Multi-Focus Group Method in Systems Analysis and Design: A Case Study', *PLoS One*, vol. 8, no. 3, pp. 1-16.
Wu, R.M.X., Zhang, Z.W., Yan, W.J., Fan, J.F., Gou, J.W., Liu, B., Gide, E., Soar, J., Shen, B., Fazal-e-Hasan, S., Liu, Z.Q., Zhang, P., Wang, P.L., Cui, X.X., Peng, Z.F. & Wang, Y. 2022, 'A Comparative Analysis of The PCA and Entropy Weight Methods to Establish the Indexing Measurement', *PLoS One*, vol. 17, no. 1, pp. 1-26.
Wu, R.M.X., Yan, W.J., Zhang, Z.W., Guo, J.W., Fan, J.F., Liu, B., Shi, Y., Shen, B., Zhao, H.J., Ma, Y.Y., Soar, J., Sun, X.Y., Gide, E., Sun, Z.G., Wang, P.L., Cui, X.X., Wang, Y. 2021, 'A Correlational Research on Developing an Innovative Integrated Gas Warning System: A Case Study in ZhongXing, China', *Geomatics, Natural Hazards and Risk*, vol. 12, no. 1, pp. 3175-204.
Yaseen, Z.M. 2021, 'An insight into machine learning models era in simulating soil, water bodies and adsorption heavy metals: Review, challenges and solutions', *Chemosphere (Oxford)*, vol. 277, pp. 1-22. https://doi.org/10.1016/j.chemosphere.2021.130126.

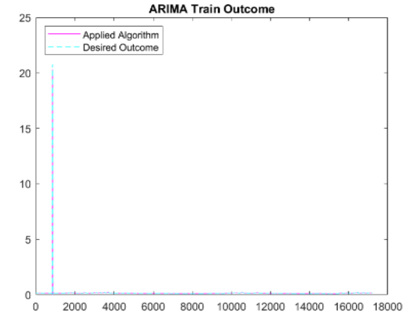
Appendix 1: ML Algorithms in China’s Indsutrial applications

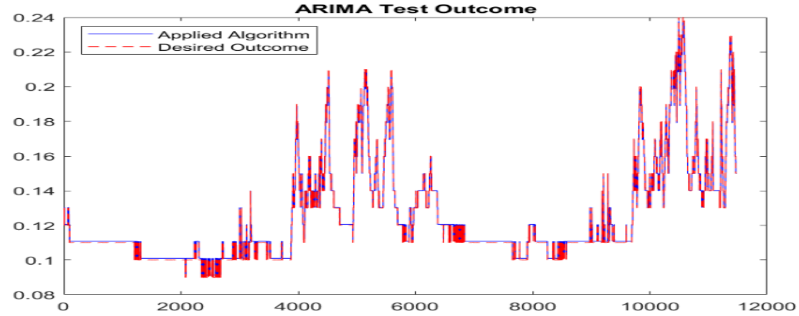
Model	2016	2017	2018	2019	2020	Total	%
BP	11	35	21	20	8	95	27.38%
SVM	23	17	19	15	11	85	24.50%
LR	3	10	11	6	0	30	8.65%
Perceptron	2	8	3	2	3	18	5.19%
RNN	3	3	5	3	3	17	4.90%
RF	0	5	3	0	5	13	3.75%
CNN	0	0	4	6	1	11	3.17%
K-means	2	1	1	1	6	11	3.17%
AdaBoost	4	1	2	2	1	10	2.88%
Bayesian Network	1	3	2	3	0	9	2.59%
KNN	1	0	1	2	3	7	2.02%
Stepwise Regression	0	3	0	2	0	5	1.44%
Naive Bayes	0	2	1	1	1	5	1.44%
SOM	0	1	1	2	0	4	1.15%
PLSR	0	2	1	1	0	4	1.15%
Logistic Regression	0	3	0	1	0	4	1.15%
LVQ	1	2	0	0	0	3	0.86%
CART	0	2	0	1	0	3	0.86%
Hierarchical Clustering	0	1	1	0	0	2	0.58%
C4.5	0	1	0	1	0	2	0.58%
RBFN	0	0	0	1	0	1	0.29%
LWL	0	1	0	0	0	1	0.29%
Projection Pursuit	0	1	0	0	0	1	0.29%
PCR	0	0	0	1	0	1	0.29%
PLS	1	0	0	0	0	1	0.29%
LDA	0	0	0	1	0	1	0.29%
GBRT	0	0	1	0	0	1	0.29%
Expectation Maximization	0	0	0	1	0	1	0.29%
Ridge Regression	0	0	1	0	0	1	0.29%
Sum	52	117	91	82	48	347	100.00%

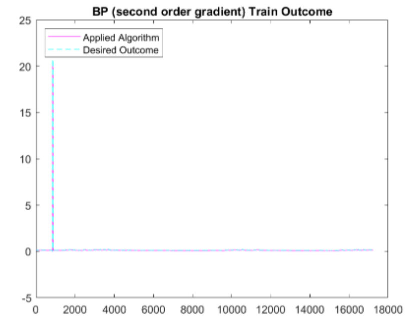
Appendix 2: Predictive Performance Assessment of ML Algorithms Between 2020 and 2023

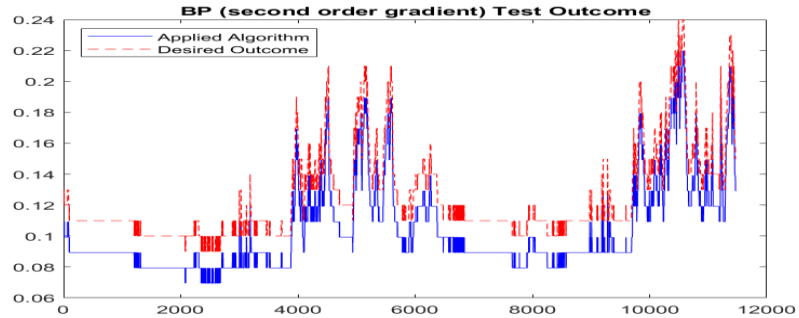
Research	AAD	AAE	AUC	CE	CoefVar	Ccross-entropy	DC	DDR	DW	EIR	GA	GI	GRC	IoA	IRR	KGE	MAD	MAE	MAPE	MBE	MdAPE	MdPE	ME	MSE	NSE	PP	OA	OOB	OE	R2	RAE	RM	RMS	RMSE	ROC	RR	SAE	SEP	Se/Sy	SMAPE	SI	SSER	sRR	TMF	Tstat
al-Swaidani et al. 2022									*									*	*											*				*					*						
Alhakamy et al. 2023																		*						*						*				*											
Ameer et al. 2022	*																													*				*				*							
Azeem 2022																			*											*															
Barashid, Munshi & Alhindi 2023																		*						*						*				*											
Carvalho 2023,				*		*												*						*													*					*			
Chung et al. 2020														*																*				*											
Das 2022																		*						*						*															
Debbarma & Ransinchung 2020		*																		*										*									*						
Elsaraiti & Merabet 2021																		*	*															*											
EL Bilali et al. 2021										*	*							*	*											*				*											
Erkaymaz 2020																		*						*						*				*											
Essam et al. 2022																		*												*		*		*											
Fang, Zhuo, Song, Yan & Zhou 2023																			*					*										*			*								
Huang et al. 2022																			*												*									*					
Huang et al. 2020																			*					*						*															
Khaliq et al. 2023			*															*																*	*										
Sui Kim et al. 2020																		*	*					*																					
Li & Song 2023																			*					*																*					
Mahmoud et al. 2022																		*	*					*										*											
Moayedi et al. 2021																		*												*	*			*											
Moghadan, Yeung & Choisne 2023																		*												*				*											
Pakzad, Roshan & Ghalehnovi 2023																		*	*	*				*						*				*							*				*
Patel et al. 2023			*		*												*									*							*			*									
Pacheco et al. 2021			*	*			*																				*	*	*					*	*										
Pakzad, Roshan & Ghalehnovi 2023																		*	*	*				*						*				*							*				*
Pu et al. 2021																		*	*					*																					
Rahbar et al. 2022								*										*							*									*											
Robeson & Willmott 2023																		*		*				*										*											
Sridharan & Shah 2023																		*																*											
Sušteršič et al. 2023																		*						*						*				*											
Sukawutthiya, Sathirapatya & Vongpaisarnsin 2021																	*													*				*											
Tabbussum & Da 2020	*																							*						*				*											
Taheri, et al. 2023																*		*							*					*				*											
Tataru-Farmus et al. 2023																														*				*											
Verhaeghe et al. 2022																		*			*	*	*							*				*											
Wang et al. 2022																		*						*										*											
Sum	2	1	3	2	1	1	1	1	1	1	1	0	0	1	0	1	2	24	13	4	1	1	1	17	2	1	1	1	1	22	2	1	1	27	2	1	2	1	2	2	2	1	0	0	2

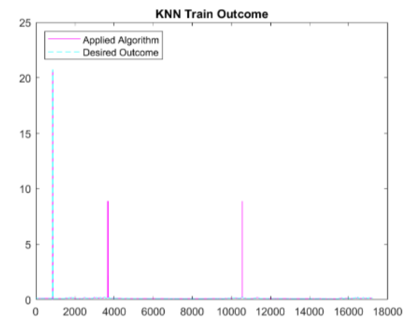
Appendix 3: Outcomes of Training and Testing

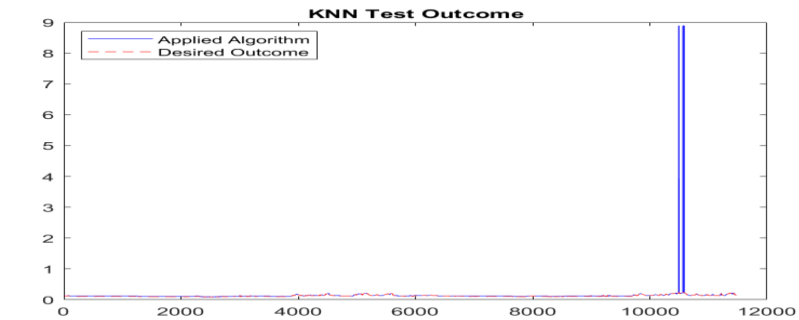


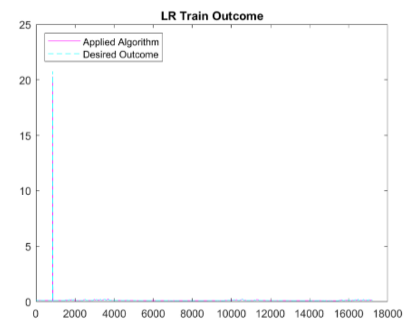


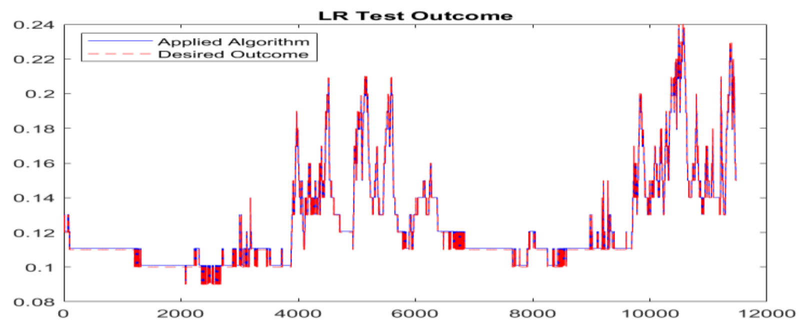


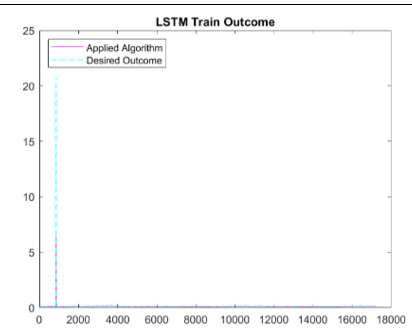


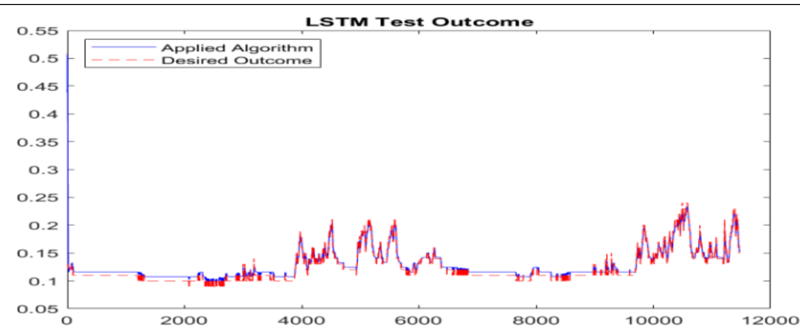


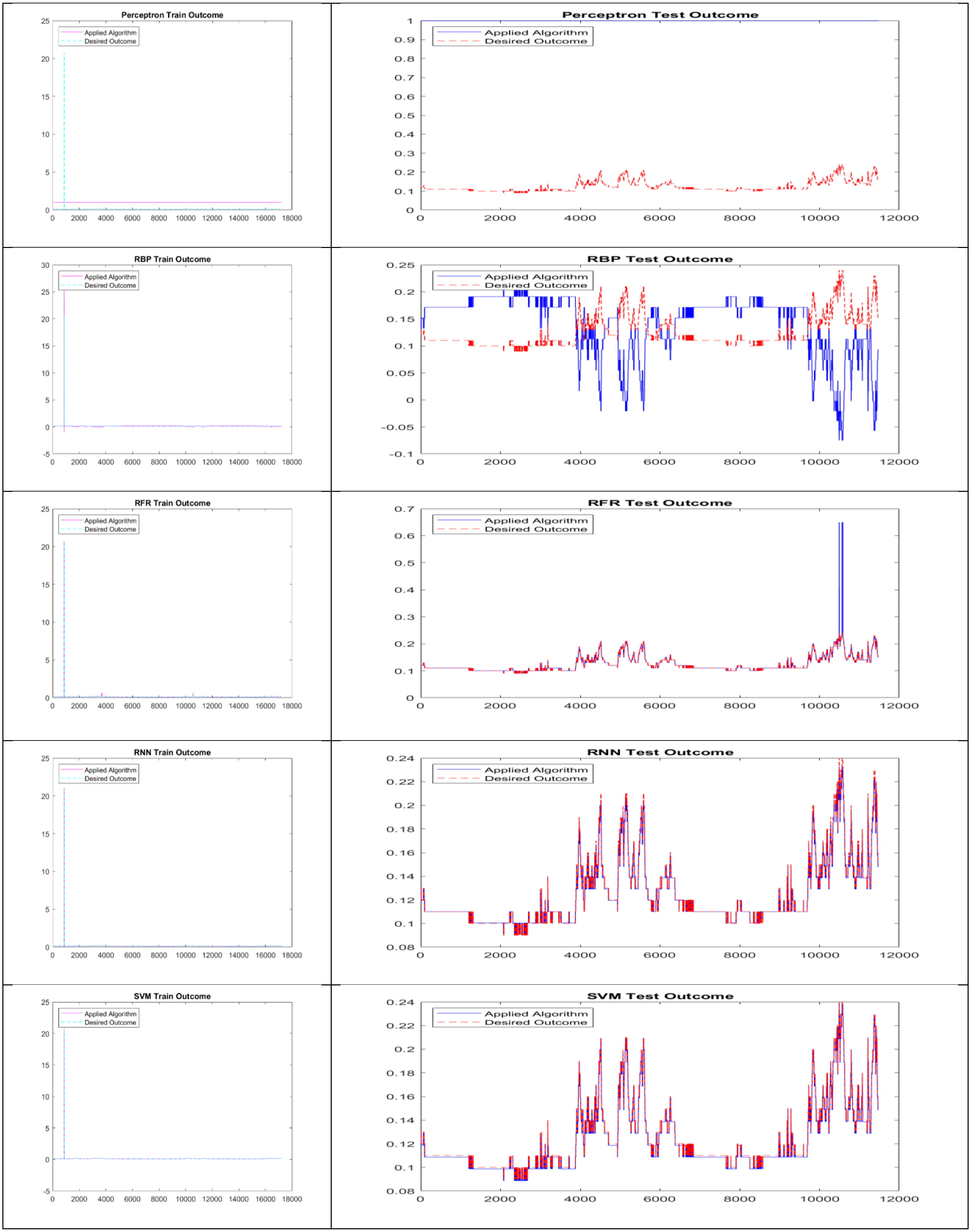











Perceptron Train Outcome
Applied Algorithm
Desired Outcome
Perceptron Test Outcome
Applied Algorithm
Desired Outcome
RBP Train Outcome
Applied Algorithm
Desired Outcome
RBP Test Outcome
Applied Algorithm
Desired Outcome
RFR Train Outcome
Applied Algorithm
Desired Outcome
RFR Test Outcome
Applied Algorithm
Desired Outcome
RNN Train Outcome
Applied Algorithm
Desired Outcome
RNN Test Outcome
Applied Algorithm
Desired Outcome
SVM Train Outcome
Applied Algorithm
Desired Outcome
SVM Test Outcome
Applied Algorithm
Desired Outcome