Journal of Information Systems Engineering and Management
2025, 10(39s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Enhanced Energy Efficiency Method in Oil and Gas Industry Using Hybrid Machine Learning Approach

[1]Ms.S. Bharathi, [2]Dr. P. Sujatha
[1]Ph.D Research Scholar School of Computer sciences Vels Institude of Science,Technology And Advanced Studies Chennai,India
karupanneer@gmail.com
[2]Professor and Head Department of information technology Vels Institute of Science, Technology and Advanced Studies Chennai, India

ARTICLE INFO

Received: 26 Dec 2024

Revised: 18 Feb 2025

Accepted: 28 Feb 2025

ABSTRACT

In the OGI, there are multiple opportunities to reduce total production energy consumption to produce the equivalent or more unrefined petroleum and flammable gas, and improvements in efficiency may be influenced by mechanical, chemical factors, what's more, other actual boundaries. The most widely recognized strategies for further developing creation energy efficiency include replacing proficient creation hardware and further developing creation processes. In addition, energy consumption forecasts can help managers scientifically plan energy use for energy generation and transition energy use to non-peak hours. However, this remains a challenging issue due to inherent complexities and vulnerabilities. Different mixes of energy utilization and apparent performance of OGI are presented to this end; our work focuses on predicting the energy utilization of OGI. First, four different prediction models Support Vector Machines, Linear Regression, Extreme Learning Machines and Artificial Neural Networks are prepared on the preparation of informational collection and then evaluated using the test data set. The second is to work on the precision of energy utilization expectations, the combination of these four models was tested, and the value of the predictor variable was taken as the normal of the outcomes of the two models. The results show that four different model predetermines energy consumption versus accuracy, as well as artificial hybrids and models of red neurons for maximum accuracy. Furthermore, hybrid models are installed in energy management systems in the OGI to oversee energy utilization in oil fields and further develop proficiency.

Keywords: Energy efficiency, oil and gas, machine learning, hybrid, models, performance metrics

I. INTRODUCTION

O&G will stay basic to worldwide monetary turn of events and success for quite a long time into the future. BP's Energy Outlook report points out that by 2040, the exclusive use of oil and natural gas will show a consistent development pattern. Additionally, global concerns about climate change have shifted the focus to the energy required to produce hydrocarbon fuels. This has driven the rise of unconventional energy sources and methods, with ongoing improvements in energy efficiency. Faced with these difficulties, the business perceives that improving energy efficiency and saving energy can make important contributions to environmental assurance and energy supply.

Meanwhile, the 2016 "Thirteenth Five-Year Plan for Public Finance and Social Development of the People's Republic of China" was unveiled. The statement also specifically suggests explicit markers for energy utilization, including the power area, sustainable power, hydro, wind, sun oriented and biomass. For example energy utilization is normal to be 6,800 to 7,200 terawatt hours, with a typical yearly development rate of 3.6% to 4.8%. In the United States, energy companies must work harder to make sound energy investments rather than produce them, Reducing energy consumption costs often offers the greatest potential for savings at the end-user level. Energy efficiency is a complex challenge for O&G companies, but it can be tackled through strategic planning, investment, and operational improvements.

Journal of Information Systems Engineering and Management
2025, 10(39s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Today, with the ever-increasing power of vast amounts of information, there are tremendous advances. With the ascent of cutting edge metering framework and the Web of Things (IOT), a lot of energy creation and utilization information are gathered and put away, and numerous scientists are Use machine learning technology. Investigate these information and dissect energy consumption behavior in light of measurable hypothesis. Therefore, smart energy forecasting methods have proven to be an effective way to improve energy efficiency in certain industrial sectors: the authors present forecasting tools with brain organizations and relapse models for predicting daily energy production in small-scale enterprises. Photovoltaic solar power generation with high prediction exactness; proposed a period series model for a single day of solar energy forecast, the model achieved maximum execution using a 9.28% error. Some researchers also explore collective methods regarding the short-term solar probability forecasts and fuzzy approaches for global solar radiation predictions. Regarding the expectation of building warming energy utilization, several artificial neural networks have also been explored: De Gouveia, S. M et. al.,(2024). Use machine learning-based models to anticipate the energy interest of indoor heating systems; conversely. The fact that some scientists straightforwardly makes it really important use regression models to calculate energy consumption. Broadly utilized relapse strategies, for example, weighted help vector relapse and different direct relapse have accomplished great outcomes. These approaches will ensure a more steady energy supply for energy companies and optimize energy demand management.

In summary, this paper addresses the question: "How do we properly anticipate the total energy required by O&G companies to produce crude O&G, and help them improve energy efficiency?" To answer this, the paper analyzes four prediction models: support vector machine, direct regression, scandal learning machine, and pseudo neural network, all of which can predict energy consumption and enhance efficiency. Additionally, a hybrid model was examined for improved performance by averaging the power outputs of the two models.

The main innovation points of this article include the following points

- Energy consumption forecasts for O&G organizations are conducted at the corporate level rather than at the oil operations level, which has proven useful for company managers and policymakers to come to informed conclusions about energy use.
- Four different forecasting methods and six hybrid models for enterprise energy consumption forecasting are discussed and analyzed.
- Four types of performance evaluation indicators are talked about to assess the forecast model. The lower the model indicator esteem, the better the model exhibition.
- The aftereffects of this paper can be used to expect the energy utilization of other O&G organizations.

II. LITERATURE SURVEY

Aditiyawarman *et al.,* (2023) explored the effectiveness of hybrid models in predicting permeability and porosity from well logs. They introduced a neuro-fuzzy system combining Artificial Neural Networks (ANN) and Fuzzy Logic (FL), leveraging the strengths of both. The model outperformed conventional methods, even without optimization.

Emeka-Okoli, S *et. al.,*(2024) can also be used to develop inventory optimization models in optimizing inventory levels by providing more accurate demand forecasts and inventory optimization models. ML algorithms can analyze large volumes of historical data, including sales data, production data, and external factors such as market trends and economic indicators, to predict future demand more accurately than traditional methods.

Khan, M. N. A *et al.,* (2024) in O&G exploring ML algorithms may prioritize certain objectives, such as maximizing production or minimizing costs, without considering broader social and environmental impacts. This can lead to conflicts of interest and negative outcomes for affected stakeholders, including local communities and the environment.

Ukato, A *et. al.,* (2024) data is collected, processed, and analyzed in the O&G industry by deploying ML algorithms at the edge, companies can perform real-time analytics, predictive maintenance, and anomaly such as ingress of corrosion detection, enabling proactive decision-making and improving operational efficiency across remote and distributed assets.

Yousefmarzi, F *et al.,* (2024) Integration of AI Technologies in Engineering Processes, in the oil and gas industry, AI technologies can be seamlessly integrated into engineering processes across the entire value chain, from

Journal of Information Systems Engineering and Management
2025, 10(39s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

exploration and production to transportation and refining. Exploration, AI-powered geophysical surveys and seismic imaging techniques enable more accurate reservoir characterization and prospect identification

Wang, Q *et al.,* (2023) Machine Learning has revolutionized demand forecasting in the oil and gas sector, enabling companies to predict future demand with unprecedented accuracy. By analyzing historical data, market trends, and other relevant factors, ML algorithms can forecast demand more effectively than traditional methods. This article explores the use of ML in demand forecasting, its benefits, and provides case studies of successful implementations

Ali, S. I *et al.,* (2023) Reservoir management plays a pivotal role in the O&G industry, encompassing the strategies and techniques employed to optimize production and recovery from subsurface reservoirs management is a multidisciplinary endeavor that requires a deep understanding of geology, fluid dynamics, engineering principles, and economic factors.

Aladwani, F *et. al.,* (2023) provides insights into the fundamental concepts of aerospace predictive maintenance, which are applicable to the oil and gas industry. Predictive maintenance is considered a subset of Condition-Based Maintenance (CBM) and leverages digital technologies for data acquisition, processing, and analysis.

III. TRADITIONAL ENERGY EFFICIENCY IMPROVEMENT METHOD

In OGI sector, there are multiple opportunities to reduce total production Energy consumption and improvements in efficiency to produce the same or greater quantities of unrefined petroleum and flammable gas can be impacted by mechanical, synthetic and other actual boundaries.

The most well-known strategies for further developing creation energy efficiency include replacing proficient creation gear and further developing creation processes. In oil refineries, for example, studies by several companies have shown that there is huge potential to further develop energy effectiveness, with key areas including utilities, radiators, process enhancement, heat exchangers, engines and engine applications.

The following cases in the literature research are very representative furthermore, generally utilized energy proficiency improvement techniques: Du et. al., (2024). By applying heater pipe gas squander heat recuperation innovation, new technologies can be used to work on the warm productivity of vacuum furnaces and meet energy-saving requirements. In request to diminish the energy utilization of pipelines, genetic algorithms are utilized to anticipate and upgrade pipelines. The exhaustive energy utilization strategy for the O&G creation process is concentrated on in the cross breed model, as shown in Figure 1; The entire O&G production process, including mechanical extraction systems and collection and transmission processes, was discussed in detail, and each significant element of the interaction chain was broke down and at long last settled. The LS-SVM model is used to anticipate the included energy consumption from oil production operations in some areas.

As well as enhancing operating conditions and equipment technology changes, changing production energy use management is likewise one of the best ways of lessening creation energy costs. This can be executed by making an association wide energy the executives plan. Many O&G companies have planned and fabricated their own energy the board frameworks with a strong commitment to energy efficiency.

Today, smart oilfields also offer new opportunities to improve energy effectiveness in O&G fields by saving working expenses. For example, in a brilliant oil field, the two specialists and designers can access data and records connected with their activities. At the point when complex issues emerge on sites that are beyond the scope of workers' knowledge, they can easily be solved by experts. Operational costs will be significantly reduced due to faster, improved decision-making. At the same time, smart oilfields can additionally give significant open doors to improve the current status of oilfield assets in supply designing and creation execution, at last lessening energy utilization in oil investigation.

This paper studies the energy utilization forecast strategy of data-driven federated machine learning models to simply improve energy efficiency and apply greater model performance to actual production at the enterprise scale rather than at oil operating stations atmosphere.

IV. PROPOSED MODEL

The proposed model consolidates three phases, as displayed in Figure 2. The objective of the primary stage is to partition the whole informational index into a few unique subsets or gatherings, considered as a grouping strategy, namely KNN. At the same time, this bunching technique is additionally reasonable for removing non-stationary or incorrect examples or perceptions from the whole informational collection. The second stage predicts energy

Journal of Information Systems Engineering and Management
2025, 10(39s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

consumption with two-dimensional contributions from crude oil petroleum and flammable gas creation in every information subset. Four different predictive machine learning methods are used at this stage, including support vector predictor machines linear regression and extreme learning. Artificial neural networks and machine learning neural networks were examined independently. The final stage aims to enforce the last forecast results through the group model. The conclusion is that ensemble models can improve the speculation capacity and exactness of AI models compared with single models. At the same time, many researchers have shown that simply combining the results of many models is beneficial for achieving better performance. In general, an ensemble model will perform better if the individual models perform well on the data set and their errors are distributed in different spaces. In our paper, we evaluate six combinations by averaging the results of the two second-stage models and then assess their performance.

V. METHODOLOGY

5.1 DATASETS

In this article, we collect two decades Energy creation and utilization information for four O&G companies. Typically, in O&G companies, this data will be announced by the energy the executive's framework, so the assortment recurrence is month to month. Company managers then adjust energy production and usage plans for the next month. We gather 12 perceptions each year, and the dataset contains a total of 960 information tests. We partition these information tests into preparing informational collections and test data sets in a ratio of 6:4. Energy utilization in the O&G industry generally includes industrial and non-industrial parts. The modern part is the immediate energy consumed in the production of oil and natural gas, while the non-industrial part mainly contributes to the daily life of oil workers, like warming, cooking, transportation, and so forth. Because of the intricacy and variety, in our work we just think about the modern piece of the non-modern part, and these data sets are represented in triple form (x, y, z), where x represents crude oil production, natural production. Flammable gas creation and modern energy utilization. However, as shown in Figure 3, the power consumption fluctuations are weakly related to time for all observations. Therefore, only oil production furthermore, gaseous petrol creation are planned as contributions for energy utilization expectation.

5.2 PREDICTION METRICS FOR MODEL PERFORMANCE

To evaluate the predictive capability of these models, which are spaced as evenly as possible, four types of metrics are introduced. First of all, forecast precision is the capacity of an indicator to predict with minimum error, which can be straightforwardly estimated by the accompanying three pointers.

$$\text{RMSE} = \frac{\sum - i(vi - vi)2}{\text{N}}$$

$$\text{MAE} = \frac{\sum_{i=1}^{N}(vi - vi)}{\text{N}}$$

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{n}\frac{(\text{vi} - \text{vi})}{\text{vi}} * 100$$

Where N represents the complete number of tests, vi is the actual value, and vi' is the anticipated esteem. Model execution is estimated by RMSE, MAE and MAPE.

5.3 CLUSTERING

The first step of the model is to categorize this large number of information tests into various subsets and afterward fabricate a particular prescient model for every information subset. From a statistical perspective, In this paper, grouping is thought of as a solo order issue with an unknown number of subsets. To do this, the following questions must be considered: What Clustering methods ought to be utilized to group these information. The quantity of subsets that thought to be made. In any case, there are many high level techniques that location both challenges, and based on the work done √√above; K-means is by all accounts the right model for this situation.

Journal of Information Systems Engineering and Management
2025, 10(39s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

As a centroid -based bunching strategy, K-means predominantly involves a two-step iterative process: first, assign each assigns the perception to its nearest subset, and then adjust the mean to minimize the distance of the total sum of squares within the group. And the process is repeated until the medium remains stationary.

In our work, the quantity of bunches must be confirmed in advance before using the K-means algorithm, what's more, there are numerous strategies to view as the ideal number of clusters by counting the quantity of bunches within and between clusters distance from the cluster or use the R-squared metric to verify clusters by assessing their Consistency. We utilize K-means to group comparative datasets into the same cluster, ensuring that the total distance within the clusters is minimized.

5.4 MACHINE LEARNING ALGORITHMS

5.4.1 Linear Regression

To understand how this ml algorithm works, imagine how you randomly place wood in order to increase their weight ranking. However, there is a catch: you cannot measure the weight of every log. You must estimate their weight based on their height and torso circumference (visual analysis) and organizing them using a combination of these visible parameters. The same goes for linear regression in machine learning.

In this process, a relationship is established between the free and subordinate factors by fitting the autonomous and subordinate factors to a line this line is called the regression line and is represented by the linear equation

$Y = a * X + b$.

In this equation:

Y – Dependent Variable

a – Slope

X – Independent variable

b – Intercept

The coefficients a and b are determined by minimizing the sum of the squared differences between the data points and the regression line.

5.4.2 Logistic Regression

Strategic relapse is utilized to gauge discrete qualities (typically twofold qualities like 0/1) from a bunch of free factors Assists in predicting the likelihood of an event by fitting the data to the logit function, commonly referred to as logistic regression.

The methods listed below are commonly employed to enhance logistic regression models:

- Include interaction terms
- Eliminate features
- Apply regularization techniques
- Utilize a non-linear model

5.4.3 Decision Tree

The decision tree algorithm in ML is one of the most well-known algorithms today; it is a supervised learning algorithm utilized for classification problems. It is applicable to both categorical and continuous dependent variables. The algorithm partitions the population into two or more uniform sets based on the most important attributes or independent variables.

5.4.4 SVM Algorithm

The SVM algorithm is a classification technique that involves plotting raw data as points in an n-dimensional space (where n represents the number of features). Each feature's values are associated with specific coordinates,

Journal of Information Systems Engineering and Management
2025, 10(39s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

facilitating the separation of data. Classifiers, or decision boundaries, are employed to divide the data and visualize it on a graph.

5.4.5 Naive Bayes Algorithm

Naive Bayes classifiers operate under the assumption that the presence of a specific feature within a class is independent of the presence of any other features. Despite the potential interdependencies among these features, the NB classifier evaluates each characteristic individually when calculating the probability of a particular outcome. NB models are not difficult to assemble and valuable for enormous datasets. It is straightforward and is known to beat even the most modern arranging techniques.

5.4.6 KNN (K- Nearest Neighbors) Algorithm

The calculation can be applied to order and relapse issues. Clearly, in the information science industry, it is involved something else for tackling arrangement issues. This is a straightforward algorithm that retains all available cases and classifies any new case by considering the majority vote of its k nearest neighbors. The new case is assigned to the class that is most similar to it, with the distance function used to determine this similarity.

KNN is easily understood by comparing it to real-world scenarios. For instance, if you want to learn about someone, it's beneficial to converse with their friends and acquaintances. Interesting points prior to choosing K Nearest Neighbor's Algorithm:

- KNN is computationally expensive
- Variables must be normalized, as higher-ranked variables can skew the algorithm.
- Information actually should be pre-handled.

5.4.7 K-Means

It is an unsupervised learning algorithm designed to address the clustering problem. A dataset is grouped into a particular number of bunches.(let's call it the number K) makes all data points in one group homogeneous and heterogeneous with those in other construct groups.

How K-means creates clusters:

- The K-means algorithm selects k points, known as centroids, for each cluster.
- Each data point is assigned to the cluster of the nearest centroid, forming the K clusters.
- New centroids are then calculated based on the current cluster members.
- Using these updated centroids, the nearest distance to each data point is determined. This process is repeated until the centroids no longer change.

5.4.8 Random Forest Algorithm

A set of choice trees is called a RF. To classify new objects based on their attributes, each tree makes a classification, and then the trees "vote" for that class. KNN is easily comprehended by relating it to real-world situations.

Each tree is planted & grown as follows:

- If the training set contains N cases, a random sample of N cases is chosen. This sample will serve as the growing tree's training set.
- If there are M input variables, provide M such that m variables are randomly chosen from M at each node, and during this process, the value of m remains unchanged. the best split at this m is utilized to split the node. During this process, the value of m remains unchanged.
- Every tree is developed to its fullest potential size. Pruning is not done.

5.4.9 Artificial Neural Network (ANN)

Back propagation involves the iterative processing of the training data set, which is used to compare the predictions of the network against each known tuple. The target values can be well-known tagged tuples for class training (in classification problems) or continuous values for predictions. For each training tuple, the weights are adjusted to

Journal of Information Systems Engineering and Management
2025, 10(39s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

minimize the mean squared error between the network's prediction and the actual target value. These adjustments are made in a "backward" direction, moving from the output layer through each hidden layer to the first hidden layer, which is why it is called back propagation. Although convergence is not guaranteed, the weights typically will eventually stabilize, at which point the learning process concludes. The steps involved are outlined in terms of inputs, outputs, and errors. However, once you become familiar with the process, each step is inherently straightforward.

5.4.10 Pseudo code for Back propagation

Input: D, A dataset comprising the training tuples and their corresponding target values l: The learning rate
Network: A **m**ultilayer feed forward network

output: A trained neural network.

Methods:

(1) Initialize all weights and biases in network;

(2) while terminating condition is not satisfied {

(3) for each training tuple X in D {

(4) // Propagate the inputs forward:

(5) for each input layer unit j {

(6) Oj = Ij; // output of an input unit is its actual input value

(7) for each hidden or output layer unit j {

(8) $I_j = \sum_i W_{ij} O_i + \theta_j$ /compute the net input of unit j with respect to the previous layer, i

(9) $_j = \frac{l}{l+e^{-i_j}}$: }// compute the output of each unit j

(10) // Backpropagate the errors:

(11) for each unit j in the output layer

(12) $Err_j = O_i \left(1 - O_j\right)(T_j - O_j)$; compute the error

(13) for each unit j in the hidden layers, from the last to the first hidden layer

(14)) $Err_j = O_i \left(1 - O_j\right) \sum_k Err_k W_{jk}$;// compute the error with respect to the next higher layer, k

(15) for each weight wi j in network {

(16) $\Delta W_{ij} = (l) Err_j O_i$ // weight increment

(17)$W_{ij} = W_{ij} = +\Delta W_{ij}$ // weight update

(18) for each bias θ j in network {

(19) $\Delta O_j = (l) Err$// bias increment

(20) $O_j = O_j + \Delta O_j$ };// bias update

(21) } }

Journal of Information Systems Engineering and Management
2025, 10(39s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

VI. HYBRID MODEL FORECASTING RESULT

Numerous energy utilization expectation issues have all the earmarks of being excessively intricate for a solitary AI model, prompting us to test a hybrid approach to investigate prediction accuracy. Often, hybrid models can compensate for the shortcomings of individual models and provide better performance. To reduce computational consumption, we simply the most delegate and usually utilized measurement, namely RMSE, was chosen instead of four measurements to assess the exhibition of these 6 blends.

Table 1: Test dataset Result of the Hybrid Model and other models

METHOD	RMSE	MAE	MAPE
SVM	2.76	0.91	5.97
LR	0.51	0.73	4.87
ELM	0.48	0.38	4.56
ANN	0.42	0.33	4.42
ANN + ELM	0.41	0.36	4.41

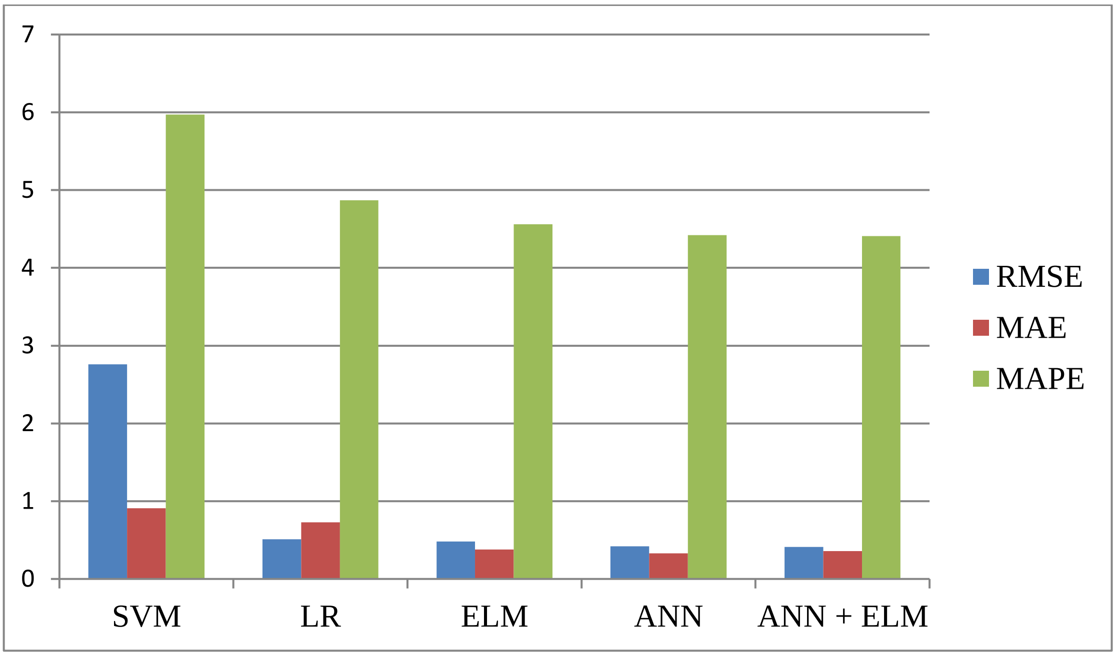

Figure 1: Performance Evaluation of Hybrid Model

As displayed in Table 1, the half breed model of ANN and ELM achieves the optimal exhibition on all datasets and the combined results are in every case better compared to any singular outcomes. Finally, we selected a hybrid model (ANN and ELM) as our last model, and the half breed model outcomes were tested on all subsets.

VII. CONCLUSION

This article first four machine learning methods commonly used for energy consumption forecasting in the O&G industry are discussed. Experimental results show that a large portion of the four dissected models can show good energy consumption prediction performance, and the coefficients of determination of these models in all data sets range from 0.70 to 0.95. For LR models, it exhibits the lowest performance for the subsequent reasons: the power

Journal of Information Systems Engineering and Management
2025, 10(39s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

utilization informational index is excessively mind boggling and has solid nonlinear connections. The SVM model also showed strong performance with polynomial kernel functions, with coefficients of determination around 0.70 for all these data sets. ANN and ELM have the most elevated exactness and have proven their viability in O&G field energy consumption prediction cases. To improve the accuracy of the last energy utilization forecast, we tested the cross breed strategy by taking the normal of its outcomes and trained on the same training dataset. The end-product shows that the cross breed model is able to attain greater prediction accuracy than individual accuracy. Among them, the half breed model of ANN and ELM showed the optimal performance and was picked as our last model for predicting the all-out energy utilization of O&G fields. We have likewise introduced this model in the energy the board course of action of the O&G industry to assist organizations with anticipating absolute energy utilization and further develop energy effectiveness.

REFERENCES

[1] Aditiyawarman, T., Kaban, A. P. S., & Soedarsono, J. W. (2023). A recent of risk-based inspection development to support service excellence in the oil and gas industry: an artificial intelligence perspective. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, *9*(1), 010801.

[2] Emeka-Okoli, S., Nwankwo, T. C., Otonnah, C. A., & Nwankwo, E. E. (2024). The evolution of CSR reporting in the oil and gas industry using machine learning and its future direction: A conceptual review. *World Journal of Advanced Research and Reviews*, *21*(3), 100-108.

[3] Khan, M. N. A., Haq, Z. U., Ullah, H., Naqvi, S. R., Ahmed, U., Zaman, M., & Amin, N. A. S. (2024). Prediction of hydrogen yield from supercritical Oil and gas industry process of sewage sludge using machine learning and particle swarm hybrid strategy. *International Journal of Hydrogen Energy*, *54*, 512-525.

[4] Ukato, A., Sofoluwe, O. O., Jambol, D. D., & Ochulor, O. J. (2024). Technical support as a catalyst for innovation and special project success in oil and gas. *International Journal of Management & Entrepreneurship Research*, *6*(5), 1498-1511.

[5] Yousefmarzi, F., Haratian, A., Mahdavi Kalatehno, J., & Keihani Kamal, M. (2024). Machine learning approaches for estimating interfacial tension between oil/gas and oil/water systems: a performance analysis. Scientific Reports, 14(1), 858.

[6] Wang, Q., Chen, D., Li, M., Li, S., Wang, F., Yang, Z.,& Yao, D. (2023). A novel method for petroleum and oil gas resource potential evaluation and prediction by support vector machines (SVM). *Applied Energy*, *351*, 121836.

[7] Ali, S. I., Lalji, S. M., Awan, Z., Qasim, M., Alshahrani, T., Khan, F., & Ashraf, A. (2023). Prediction of asphaltene stability in crude oils using machine learning algorithms. *Chemometrics and Intelligent Laboratory Systems*, *235*, 104784.

[8] Aladwani, F., & Elsharkawy, A. (2023). Improved prediction of heavy oil viscosity at various conditions utilizing various supervised machine learning regression. *Petroleum Science and Technology*, *41*(4), 406-424.

[9] De Gouveia, S. M., de Abreu Corrêa, L., Teles, D. B., Oliveira, M., & Clarke, T. G. R. (2024). Oil and Gas industry damage classification by machine learning using synthetic data. *Engineering Failure Analysis*, *156*, 107819.

[10] Du, J., Zheng, J., Liang, Y., Xu, N., Liao, Q., Wang, B., & Zhang, H. (2023). Deep pipe: Theory-guided prediction method based automatic machine learning for maximum pitting corrosion depth of oil and gas industry. *Chemical Engineering Science*, *278*, 118927.

[11] Gao, M., Liu, Z., Qian, S., Liu, W., Li, W., Yin, H., & Cao, J. (2023). Machine-Learning-Based Approach to Optimize CO2-WAG Flooding in Low Permeability Oil Reservoirs. *Energies*, *16*(17), 6149.

[12] Odimarha, A. C., Ayodeji, S. A., & Abaku, E. A. (2024). Machine learning's influence on supply chain and logistics optimization in the oil and gas sector: a comprehensive analysis. *Computer Science & IT Research Journal*, *5*(3), 725-740.

[13] Peng, J., Wen, L., Mu, X., & Xiao, J. (2023). The evolving centres of gravity in China's oil and gas industry: Evidence from infrared radiation imaging gas flaring data. *Energy for Sustainable Development*, *73*, 263-279.

[14] Shaik, N. B., Jongkittinarukorn, K., Benjapolakul, W., & Bingi, K. (2024). A novel neural network-based framework to estimate oil and gas Industry life with missing input parameters. *Scientific Reports*, *14*(1), 4511.

[15] Qin, G., Xia, A., Lu, H., Wang, Y., Li, R., & Wang, C. (2023). A hybrid machine learning model for predicting crater width formed by explosions of oil and gas industry. *Journal of Loss Prevention in the Process Industries*, *82*, 104994.

[16] Li, J., Guo, Y., Fu, Z., Zhang, X., & Shen, F. (2023). An intelligent energy management information system with machine learning algorithms in oil and gas industry. *Wireless Communications and Mobile Computing*, *2023*(1), 3385453.

[17] Ukato, A., Sofoluwe, O. O., Jambol, D. D., & Ochulor, O. J. (2024). Technical support as a catalyst for innovation and special project success in oil and gas. *International Journal of Management & Entrepreneurship Research*, *6*(5), 1498-1511.

[18] Masudin, I., Tsamarah, N., Restuputri, D. P., Trireksani, T., & Djajadikerta, H. G. (2024). The impact of safety climate on human-technology interaction and sustainable development: Evidence from Indonesian oil and gas industry. *Journal of cleaner production*, *434*, 140211.

[19] Lin, X., Li, G., Wang, Y., Zeng, K., Yang, W., & Wang, F. (2024). Machine learning in Intelligent Identification of Fiber-Optic Vibration Signals in Oil and Gas industry. Journal of Pipeline Science and Engineering, 100184.

[20] Hussain, M., Alamri, A., Zhang, T., & Jamil, I. (2024). Application of Artificial Intelligence in the Oil and Gas Industry. In *Engineering Applications of Artificial Intelligence* (pp. 341-373). Cham: Springer Nature Switzerland.

[21] Goliatt, L., Saporetti, C. M., Oliveira, L. C., & Pereira, E. (2024). Performance of evolutionary optimized machine learning for modeling total organic carbon in core samples of shale oil and gas fields. *Petroleum*, *10*(1), 150-164.

[22] Arinze, C. A., Izionworu, V. O., Isong, D., Daudu, C. D., & Adefemi, A. (2024). Integrating artificial intelligence into engineering processes for improved efficiency and safety in oil and gas operations. *Open Access Research Journal of Engineering and Technology*, *6*(1), 39-51.

[23] Bigdeli, A., & Delshad, M. (2024). The Evolving Landscape of Oil and Gas Chemicals: Convergence of Artificial Intelligence and Chemical-Enhanced Oil Recovery in the Energy Transition toward Sustainable Energy Systems and Net-Zero Emissions. *Journal of Data Science and Intelligent Systems*, *2*(2), 65-78.

[24] Xu, D., Zhang, Z., He, Z., & Wang, S. (2023). Machine learning-driven prediction and optimization of monoaromatic oil production from catalytic co-pyrolysis of biomass and plastic wastes. *Fuel*, *350*, 128819.

[25] Wang, M., Hui, G., Pang, Y., Wang, S., & Chen, S. (2023). Optimization of machine learning approaches for shale gas production forecast. *Geoenergy Science and Engineering*, *226*, 211719