International Journal of Research in Engineering, Science and Management
Volume 7, Issue 3, March 2024
https://www.ijresm.com | ISSN (Online): 2581-5792

Exploring the Landscape: Unveiling Dominant Approaches, Benefits and Potential Pitfalls in Utilizing Diverse Machine Learning Methods for Facies Classification – A Review

Ofoh Ifeyinwa Juliana[1*], Onyekuru Samuel Okechukwu[2], Ikoro Diugo[3], Opara Alexander Iheanyichukwu[4]
[1]*Department of Geology, Federal University of Technology, Owerri, Imo State, Nigeria*
[2,4]*Professor, Department of Geology, Federal University of Technology, Owerri, Imo State, Nigeria*
[3]*Associate Professor, Department of Geology, Federal University of Technology, Owerri, Imo State, Nigeria*

***Abstract*: Facies classification in geosciences has witnessed a remarkable transformation with the integration of diverse machine learning (ML) methods. An overview of the dominant approaches, benefits and potential pitfalls encountered in the application of different ML techniques for facies classification is presented in this review paper. ML technologies like convolutional Neural Networks (CNNs), Support Vector Machines (SVMs), Random Forests (RFs), and Deep Learning architectures have revolutionized the way geoscientists interpret subsurface reservoirs using well logs, seismic data, and core samples. The benefits of deploying Machine Learning techniques in facies classification range from enhanced speed and accuracy of interpretation, through facilitating the extraction of valuable geological insights from complex datasets, to flexible handling of multi-modal data, thus allowing for the combining many data sources to improve classification accuracy and enhance informed decision-making in exploration and development projects. However, the use of ML methods in facies classification are inherent to potential pitfalls and significant challenges, some of which include sparse data availability and poor data quality. Robust model training necessitates large, labeled datasets that are often costly and time-consuming to curate. In addition, model interpretability remains a major concern, since the 'black-box' nature of some ML algorithms can hinder geoscientists' ability to understand and trust the results. Overfitting, model generalization issues, and the risk of biases in training data are additional problems that must be addressed.**

***Keywords*: Artificial Neural Network, Core data, Facies classification, Geosciences, Machine Learning, Reinforcement learning, Seismic data, Supervised learning, Well log data, Unsupervised learning.**

1. Introduction

The term "Facies" can be defined as a volume of rock with certain properties, that might be any visible feature of rocks (including their general makeup, look, or formation state) and any variations in those features across a given region (Reading, 1996). Additionally, it comprises most of a rock's properties, such as its physical, chemical, and biological makeup that sets it apart from nearby rocks. (Parker, 1984). Facies (or lithofacies) classification is the procedure of designating a kind or category of rock to a given sample inspired by the observed characteristics. It has become crucial to classify diverse lithofacies in well logs and seismic interpretation since different rock types have varying permeabilities and fluid saturation levels for a given porosity.

Recent advancements in machine learning (ML) technologies have grown in favor within the petroleum sector where it is used as an invaluable tool for informed business decision-making (Xu et al., 2022). The recent growing interests to adopt ML approaches in geoscience interpretations has been spurred by the expansion of so-called big data and the increase in computing capacity. Within the vanguard of artificial intelligence (AI) technological practice is machine learning, a term that describes essentially a set of data analysis techniques including classification, regression, and clustering. (Hall, 2016). Some of the ML applications that has been presented to the community of geoscientists include support vector machines (SVMs), random forests (RFs), and artificial neural networks (ANNs) (Kuwatani et al., 2014; Bolandi et al., 2017; Wrona et al., 2018; Ai et al., 2019; Bolton et al., 2020; Lee et al., 2022). The difficulties associated with traditional manual interpretation appear to be amenable to automation of lithofacies categorization using the ML technologies, which have proven to its capabilities in enhancing and supplementing human analysis. Recent applications of the ML techniques in reservoir characterization include lithofacies classification (Leila et al., 2013; Paolo et al., 2017), depositional facies forecasting (Randell et al., 2019; Jing et al., 2022), well log correlation (Hiren et al., 2018; Tran et al., 2020), seismic facies classification (Satinder et al., 2018; Seth et al., 2019), and seismic facies analysis (Thilo et al., 2018; Vladimir et al., 2022), among other tasks.

The present paper presents a critical review of the dominant approaches, benefits and potential pitfalls in utilizing diverse machine learning methods for facies classification.

*Corresponding author: ifeyinwaofoh@gmail.com

A. Rational of the Review

This review was done to comprehensively explore and analyze various machine learning methods employed in the field of facies classification by thorough examination of dominant approaches, highlighting their benefits and drawbacks. The goal is to provide a nuanced understanding of the diverse techniques used for facies classification, enabling researchers, practitioners, and stakeholders to make informed decisions when selecting and implementing machine learning methods in geological and petrophysical applications.

B. Objective of the Review

The objective is to comprehensively investigate, analyze, and give a holistic comprehension of the utilization of diverse approaches for machine learning for facies classification in geoscience and reservoir engineering by: Identifying Dominant Approaches, evaluating benefits, assessing potential pitfalls, showcasing real-world applications and providing future directions.

C. The Role of Machine Learning in Enhancing Facies Classification Accuracy

Mathematical models have been utilized for lithology identification from the era of well logs (Definer et al., 1987). Following then, utilizing a computer program to carry out the identical task has been essential for good log analyzers, especially when it comes to cutting processing time while enhancing accuracy. But up until recently, the main foundation of these prediction models was fixed empirical connections, which were frequently restricted to a small number of depositional settings and geographic regions. However, models which represent non-linear relationships can be produced by new automated procedures based on machine learning techniques. for larger observational density and a wider variety of contexts. These models, which were initially put out using neural networks (Baldwin et al., 1990; Rogers et al., 1992), can yield findings with improved precision, that can then be utilized to increase the clarity of reservoir models employed for geophysical interpretation. Facies classification can be used, for example, in seismic interpretation to restrict important properties like permeability, fluid saturation, and porosity. In order to define the facies consistency and dispersion among data points for stratigraphy, classifications might be applied (Ngui et al., 2018). A lot of study has recently been done on the use of ML techniques for classifying facies, encompassing execution evaluation of various learning algorithms used for this job (Xie et al., 2018).

Nevertheless, an array of untapped research avenues is now more accessible than ever thanks to the numerous machine-available techniques and the capacity to combine these into arbitrary intricate patterns. This is because the cost of geophysical interpretation software and data processing services keeps coming down, making testing more accessible. In order to measure and enhance prediction accuracy, a large portion of the current research on the subject has concentrated on using one algorithm, assessing many methods individually, or integrating homogenous algorithms into ensembles (Gifford & Agah, 2010). An alternative strategy uses a mix of learners (or agents) to enhance additional variables with optimizing prediction accuracy.

2. Overview and Basics of Machine Learning Approaches

A. Overview of Machine Learning Methods

In accordance with the types of data it uses, machine learning techniques may be categorized into four main groups: (1) supervised learners, who use input and output data to build a model for prediction (2) unsupervised learners who solely use inputs to organize and understand data (3) semi-supervised and (4) reinforcement learning. These can also be separated into groups according on if the outcome of the model produces distinct outputs or continuous signals.

1) Supervised Learning

Supervised machine learning techniques employ a training dataset, or known output, to create predictions. These algorithms may be split into two broad categories: regression algorithms and classification algorithms. Although there are differences in implementation, all classifiers fundamentally execute two tasks. Firstly, they iteratively forecast using the training dataset as input, and secondly, they correct using the genuine worth that is there in the training set. Decision Trees, Random Forests (RF), Support Vector Machines (SVM), Linear and Logistic Regression, Naive Bayes, Linear Discriminant Analysis, k-Nearest Neighbor (KNN) methods, and Neural Networks are types of commonly used supervised machine learning techniques. Other examples of supervised ML include; epileptic seizure detection can provide specific to patients' detectors that are very accurate in preventing bodily harm and death while promptly identifying seizure onsets (Kharbouch et al., 2011).

2) Unsupervised Learning

The organization of a dataset is automatically determined through strategies for unsupervised learning, namely additionally referred to as methods for data-driven exploration in data figures. This reduces relying on preconceived classification or another external limitations and presumptions. Due to the fact that unsupervised approaches determine which manifold is most fit with a collection of data that is just on the available input data, they are thought to be more robust, or resilient to the dataset's severe findings or anomalies. In unsupervised machine learning, without human intervention, unsupervised learning examines unlabeled datasets. By giving matching labels, the algorithm splits the samples inside many partitions based only on the training data's features. Autoencoders and principal component analysis constitute two popular instances of unsupervised approaches. Unsupervised learning is typically employed for automatically identifying a user's friends on social networking sites like Facebook or Google, or for determining the greatest quantity of mails sent to a specific person and categorized collectively.

3) Semi-supervised Learning

One way to think of semi-supervised learning is as a combination of supervised and unsupervised techniques as described previously since it incorporates data that is both labeled and unlabeled (Zhou & Belkin, 2014). The primary objective of a semi-supervised learning framework is to provide predictions that outperform those produced with simply the labeled data from the model. A lot of applications for this technique include text categorization, fraud detection, data labeling, and machine translation. Reinforcement learning, a situation-driven approach, is predicated on several algorithms that are often executed in a sequential fashion to automatically determine the optimal course of action in a specific setting in order to maximize efficacy (Buşoniu et al., 2010). It is therefore a useful tool for creating AI models that might improve the efficiency of intricate systems like supply networks, manufacturing, robotics, and autonomous vehicles.

4) Reinforcement Learning

The foundation of reinforcement learning is a series of algorithms that, in most cases, work in succession to automatically assess the best conduct in a given setting and increase its effectiveness; this is known as an ambient-driven approach (Busoniu et al., 2010) in every step, a reinforcement algorithm—also called a "agent"—acts and makes predictions about the characteristics in a subsequent step depending on the forecast and bases the upcoming step on the characteristics at the current and prior phases.. Based on the prediction, incentive or penalties is applied.

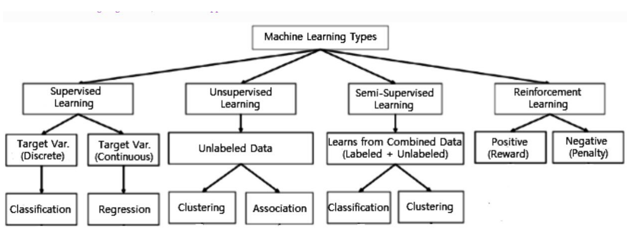

Fig. 1. Different kinds of machine learning methods after (Sarker, 2021)

B. Basic Machine Learning Approaches Commonly Used for Facies Classification

1) Decision Trees (DT)

A Decision Trees can execute both classification and regression tasks. Boolean tests of one or more properties are the foundation of the DT algorithm. The model is made up of child nodes, actions connected to each node, and a node called the root, frequently referred to as the node that serves as the parent. A graphical diagram or chart is provided which enables users to visually determine the potential outcome of any decision made. The entropy, classification error, or Gini index are used to determine which characteristic is the best, and can be expressed using the equations below in equ (1) and (2) (Sarker, 2021).

$$GINI = \sum_{m=1}^{k} p_{ik}\,(1 - p_{ik}) \tag{1}$$

$$Entropy = -\sum_{m=1}^{k} p_{ik}\, \log\, p_{ik} \tag{2}$$

Imagine a Yes-No DT created to help people decide choosing between A and B, that is, either to act on A or B (Fig. 1) below, the primary objective of DT is to constantly divide the data until it enters its particular group or reaches a predetermined depth in the tree.

The output of the method is unstable if the training data changes, and the quantity of observations from each class may have an impact. The Random Forest (RF) approach, which is a collection of several trees from the same dataset developing concurrently, can be used to get around this restriction. After a majority vote among its trees, the RF algorithm chooses the group receiving the most votes as the model's forecast. The effectiveness of DT and RF models has been examined by (Maucec & Garni, 2019).

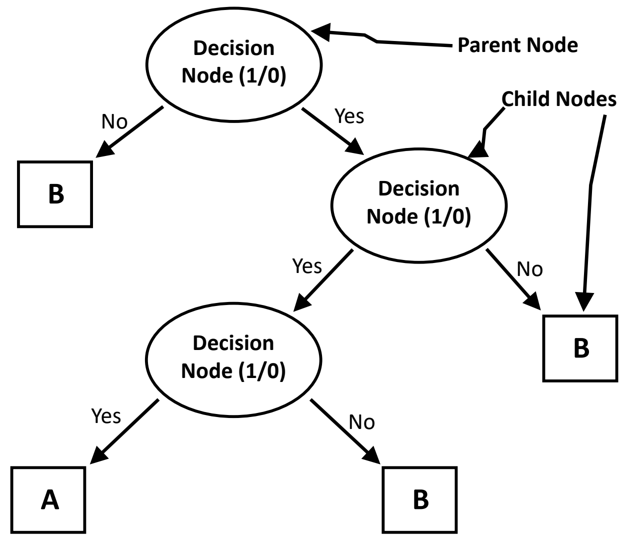

Fig. 2. A DT for deciding between options A and B, yes-no (Sachin & Karmakar, 2020)

2) Random Forest (RF)

Using a collection of graphs known as decision trees, Random Forests are ensemble supervised machine learning algorithms that model a dataset and subsequently utilize that model to generate predictions. Tree Baggers and Random Decision Forests are other names for this procedure (Breiman, 2001). Bagging (i.e., short for Bootstrap Aggregating) refers to the process of model averaging, which links the outcomes of randomly chosen training data sets in order to enhance categorization. The subtlety of the A single decision tree may have a tendency to overfit a given training dataset, showing low bias but a significant variance. However, the Random Forest algorithm can drastically decrease variance while taking advantage of a relatively small possible boost in bias by averaging several randomly selected decision trees (or models) from a large set, each of which was trained using a random subset of the training data. Additionally, the Random Forest method offers the chance to look at the model that was created during training, which is useful when trying to figure out why a certain sample was classified in a certain way.

3) Support Vector Machines (SVM)

SVM mostly performs classification problems, but it may additionally carry out regression tasks using the title Support Vector Regression (Chamkalani et al., 2013). We must train the algorithm with (2) two clearly distinct class names in order to do classification The method chooses a linear or non-linear hyperplane which is most accurate in classifying fresh information by determining which side of the hyperplane dividing the two predetermined groups it will belong to. The weight value of the hyperplane determines its direction, but the biasedness measure determines how far it deviates from the origin. Key elements or pillars of support vectors are the data points in an SVM with linear bounds that have the greatest influence on classification. For simplicity, we create a mathematical function called kernels to restructure the data points in the case of a non-linear hyperplane such that a linear hyperplane may be used to classify them linearly.

In several real-world applications, (Li et al., 2000; Lu et al., 2001; Choisy & Belaid, 2001; Gao et al., 2001; Kim et al., 2001; Ma et al., 2001; Van Gestel et al., 2001; Al-Anazi & Gates, 2009a, 2009b, 2009c), support vector machines (SVM) have recently demonstrated high generalization (prediction) efficiency.

4) Artificial Neural Networks

In the early 1980s, Artificial Neural Networks (ANNs) gained a lot of interest as a model of biologically motivated intelligence. as the neuroscience sector made significant strides in its application, sparking a keen interest in comprehending the significance of neural network models (Mohaghegh et al., 1994). Large collections of algorithms that represent artificial neural networks are able to establish relationships among extremely unusual nonlinear variables and produce advanced, precise, and dependable solutions to challenging problems by learning from experience (Ali, 1994).

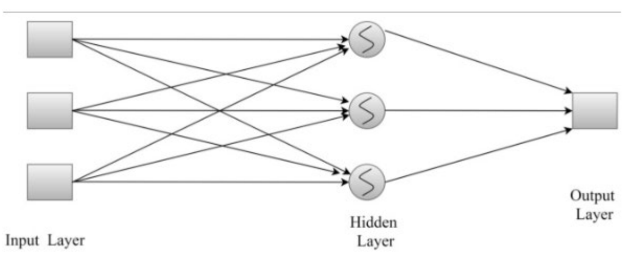

Fig. 3. An illustration of a basic neural network (after Anirbid, 2021)

Computer systems called neural networks make an effort to mimic the activities of the brain. Nonlinear data modeling methods like ANNs are made to handle complicated tasks with a lot of inputs. It comprises of an entry level, an outcome level, and a few concealed layers. The input layer gives each input its due weight. Both the input layer and the hidden layer have corresponding bias values. The hidden layer's job is to execute the weighted inputs' summing to determine the value for the layer that follows. In order to determine the value of the output(s) in the output layer, it uses an activation function. A significantly more sophisticated neural network technique that is frequently utilized in research is the probabilistic neural network (PNN). It may also be integrated with a variety of different algorithms, such as the Neural-Decision Tree (NDT) created by (Li et al., 2013). ANN is a powerful machine learning technique for handling challenging issues. ANN is most frequently utilized in the oil and gas industry to tackle complicated, nonlinear issues that cannot be resolved by linear relationships. Information, including hidden neurons, is sent forward using feed forward-ANN (FF-ANN) (Ashena & Thonhauser, 2015). Neural networks may be used in the petroleum sector for seismic pattern recognition, drill bit diagnostic, gas well yield enhancement, sandstone lithofacies detection, well performance prediction, and optimization (Ali, 1994).

5) k-Nearest Neighbors (KNN)

The k-nearest neighbors (KNN) method is an easy and straightforward technique that has grown in popularity. Given that the goal variable is present in the dataset, it falls under the category of supervised machine learning. Both classification and regression issues are addressed with it. The number of neighbors to evaluate in order to forecast a class or an output value estimation is represented by the letter k in the KNN method. KNN allocates the majority of its neighbors' categories while solving a classification problem, but it uses the mean of their neighbors' target variables when solving a regression problem.

6) Logistic Regression

The most-basic machine-learning model for categorization issues is logistic regression, which is similar to linear regression. Logistic regression and other supervised machine learning methods can be used if the labels are given while training. Instead of producing the actual result, the logistic regression model first generates the Airbed input attributes' weighted sum (as well as a bias term) (Géron, 2017).

7) Principal component analysis (PCA)

The method of principal component analysis is employed in performance forecasting that leverages common patterns and trends from vast amounts of information. The principal components technique is typically employed to predict output from shale reservoirs rich in liquid. Using Singular Value Decomposition (SVD), the primary component was computed. These computed primary components were utilized by (Makinde & Lee, 2019) to anticipate oil output. The model proved helpful in predicting output with a respectable level of accuracy. Channelized reservoirs were mapped using PCA based on the Cumulative Distribution Function (also known as CDF-PCA). According to (Chen et al., 2014), their findings demonstrated the improved and consistency of the geological facies, reservoir characteristics, and production projection model using CDF-PCA. China's natural gas business was evaluated for sustainability using principal component analysis. Employing PCA, the natural gas sustainability index was found and assessed. According to the findings, sustainability increased between 2008 and 2013 as a result of rising supply and demand (Dong et al., 2015).

3. Benefits of Machine Learning in Facies Classification

A. Machine Learning Solution

ML has swiftly gained popularity in petrophysics and has shown to be an effective solution to a variety of issues. Notwithstanding the enormous amount of enthusiasm that has been sparked by it in several distinct fields, ML is not a magic solution to every issue. Because of how serious the problem is or how much high-quality data is available, ML solutions may not be the best options. We may wish to assess if ML is crucial or cost-effective before applying it to any petrophysical problem. Finding a solution to help with corporate decision making is the goal of employing machine learning, regardless of the model or technique employed.

However, from a physics standpoint, problems cannot be addressed if a physical model is lacking or too complicated. As a technology driven tool, ML can often offer a solution in this situation with quantifiable unreliability (Khan et al., 2018; Basu et al., 2020; Chen et al., 2022). All models are incorrect, as stated, yet some are really helpful. It is helpful if the ML solution is shown to be effective in assisting in corporate decision-making.

B. Automated Machine Learning for Consistency and Productivity

The higher productivity that robots provide over humans in jobs such data rectification and quality assurance, data labeling, grouping as well as well log correlation is another significant benefit of employing ML (Brazell et al., 2019; Bakdi et al., 2020; Liu et al., 2021). Unlike human interpretation, which is dependent on information, experience, and abilities, machine work relies on mathematically based models. Since ML produces more reliable outcomes, human-produced goods may be subjective and prejudiced. As a result, machines can complete some monotonous tasks far more quickly and consistently than humans.

C. High-dimensional Data Analysis

High-dimensional data poses a challenge to human recognition capabilities, including lists, pictures, patterns of waves, and 3D volumes. Accurately labeling and interpreting highly dimensional data is still challenging for humans, even with the aid of contemporary 3D visualization. Pinpointing various minerals on a narrow slice with great magnification picture or tracking each pore on a rock's 3D volume CT scan is almost difficult. To solve these problems, we must turn to technology such as ML to seek solutions.

4. Limitations of Machine Learning in Facies Classification

A. Paucity of Accurate and Representative Data

Despite data being correct, they may not be representative (Ma & Amabeoku, 2015). Numerous factors, including as tool image quality, testing, and sample capture, and physical sample modification prior to testing, might lead to data that is not representative. If the data used is not indicative of the desired issue, the results from both ML and human effort would be unreliable. With the help of subject matter experts, ML approaches may be used to find outlier data (Akkurt et al., 2018; Misra et al., 2019) and rectify it. However, it is not anticipated that a machine would be able to notice this data quality issue, which will have an impact on the predicted outcomes, if ML modeling uses erroneous unrepresentative data.

B. Inconsistency in Data

Another problem that combining ML and real-world modeling approaches encounter is data consistency. Between fields and between wells, there may be major differences in the data gathering parameters. The instruments utilized for data collecting may differ from one service provider to another, since technology continues to advance quickly over time. With improvements in drilling technology and drilling fluids, borehole conditions might alter. Vertical and lateral variations in geology are constant. Modifications and fluid transformations (gas, oil, and water) are occurring in reservoirs in geomechanics with field development due to pore pressure variations. Although complex, these adjustments must be made in order to assure the consistency and quality of data from many wells or sites.

C. Size, Quality and Relevancy of the Data

Most machine learning (ML) techniques require "Big data" to educate the model. The adage "the bigger the better" could be accurate when data is accurate and representative. However, in the field of petrophysics, the data collected are frequently sparse; evaluations on a hundred of core plugs for a specific rock parameter are usual. Data quality drastically declines as data sets get larger due to the possibility that the measurements made for each set may not follow the same process, as was previously addressed in the section on data consistency machine learning, model relevance, generalization, and data labeling efficiency are still dependent on statistics. to mention but a few are some of the few limitations of ML.

5. Case Studies and Applications

Real-world applications of machine learning in facies classification with successful examples on petrophysical issues, including rock type and facies categorization which were inherently solved by machine learning techniques (Hall, 2016; Zhang & Zhan, 2017, Tahiru et al., 2021; Anirbid et al., 2021) and incomplete log prediction (Akinnikawe et al., 2018; Singh et al., 2020; Tokpanov et al., 2020). The majority of publications provide effective examples of applying one or more machine learning methods to tackle difficult petrophysical issues utilizing a variety of data sources. Few publications, nevertheless, address the restrictions associated with employing ML and the prerequisites needed to guarantee successful applications.

Below is a literature review of some authors, study aim, approach and findings in Table 1.

6. Future Trends and Directions

The past few decades have witnessed a notable advancement in artificial intelligence (AI), and machine learning (ML) in specific, as vital tools for intelligently analyzing such data and creating the matching real-world applications (Koteluk et al., 2021; Sarker, 2021b). ML is influencing developments in technology and how they are used in the real world on a global scale. Google Trends indicates that interest in "AI" and "ML" has grown significantly during the previous five years. Significant information about worldwide access to AI and ML across countries (i.e., Italy, China, the USA, Israel, the UK, and the Middle East) may be found by a Google search even if we

Table 1
Literature review

Reference	Study aim	Approach	Findings
Alexander et al. (2019)	The study aimed at improving facies classification for geological modeling by utilizing Machine learning algorithms that were selected based on dataset	Supervised machine learning algorithm. Ensemble of Decision Trees algorithm associated with gradient boosting	Proposed a supervised machine learning algorithm for facies classification and developed a workflow to enhance the dataset for classification.
Julian Thorne et al. (2019)	Machine Learning was applied in identifying reservoir facies using petrophysical and geophysical data points	Obtaining seismic data points of petrophysical and geophysical parameters. Identifying correlated clusters of petrophysical parameters	The study Identified correlated clusters of petrophysical parameters. Generation of multi-dimensional clusters of seismic data points
Vijesh Chandra et al. (2022)	The authors applied feature augmentation machine learning models for facies identification using well log and multi-scale image data from whole cores, which can be extended to multiple wells.	Feature augmentation machine learning models - Integration of well log and multi-scale image data	Accuracy of the model improved by up to 80% compared to using conventional well data alone. Incorporation of digitally derived data from whole core CT and thin section petrographs improved the accuracy of the model.
Asedegbega et al. (2021)	The purpose of the study was to use Machine Learning for reservoir facies classification.	Support vector machine, random forest, decision tree, extra tree, neural network (multilayer preceptor), k-nearest neighbor and logistic regression model were used. Jaccard index and F-1 score were for evaluation	Support vector machine: Jaccard index - 0.73, F1-score - 0.82 - K-nearest neighbor: Jaccard index - 0.91, F1-score - 0.95
Camila Martins Saporetti et al. (2021)	The research applied the unsupervised extreme learning machine (US-ELM) to cluster petrographic data collected from the Parana Basin, Brazil, and used principal component analysis (PCA) to remove redundant attributes.	unsupervised Extreme Learning Machine (US-ELM) and Principal Component Analysis (PCA) were used for evaluation.	Hybrid US-ELM outperformed methods commonly used in the literature. Higher average results for accuracy, silhouette metrics, and adjusted rand score.
Marco Ippolito et al. (2021)	The research aimed at using multi-agent approach to reduce bias introduced during training and provide a basis for producing a probability distribution for every sample as opposed to a categorization that is discrete and represents lithological regime continuity.	Unsupervised and supervised machine learning approach was used.	Facies classification from well logs is important in reservoir characterization. - Machine learning methods can improve classification accuracy.
Mehran Rahimi and Mohammad Ali Riahi (2022).	The article focused on reservoir facies classification using random forest and geostatistics	Reservoir facies classification was done using random forest and geostatistics	Reservoir facies model estimated with high accuracy (95%) - APE value of sequential indicator simulation model is less than indicator kriging model.
Jing-Jing Liu and Jianchao Liu, (2022)	The research aimed at a novel hybrid deep learning model based on the efficient data feature-extraction ability of convolutional neural networks (CNN) and the superior capacity of long short-term memory networks (LSTM) to characterize time-dependent characteristics in order to carry out lithological facies-classification tests.	Hybrid CNN-LSTM model for lithofacies classification - Borderline synthetic minority oversampling technique (BSMOTE) for data balance	Hybrid CNN-LSTM model achieved 87.3% accuracy - Processed data balance improved lithofacies classification accuracy

are aware that such statistics do not fully depict the situation. Fields addressing data-intensive issues have also been greatly impacted by machine learning including supply chain management, consumer services, and identifying errors in complicated systems. Similar wide-ranging impacts have been observed across disciplines, as ML techniques helped researchers classify cancer using DNA microarray analyses (Tan & Gilbert, 2003; Wang et al., 2005).

Generally speaking, the kind and efficiency of a machine learning solution is determined by both the data and the learning algorithms' performance. In actuality, researchers are currently now starting to investigate the potential of ML algorithms for studying systems that improve with usage, despite the surge in fascination with these fields over the previous ten years.

A. Emerging Trends in Machine Learning and their Applications

To make our jobs easier, machine learning models may be applied in many different sectors. The following are discussion of machine learning uses and future trends.

B. Analytical Prediction and Thoughtful Decision-Making

Data-driven predictive analytics for intelligent decision-making is a key machine learning application field (Cao, 2017; Mahdavinejad et al., 2018). In order to anticipate an unknown result, predictive analytics relies on identifying and taking application of correlations amongst anticipated parameters from historical data and explanatory factors (Han & Kamber 2011). For example, pinpointing offenders or culprits after a crime is done, or spotting credit card fraud while it occurs. Another use case is the improvement of inventory management, prevention of out-of-stock scenarios, and e-commerce logistics and warehousing through the use of machine learning algorithms by merchants. Numerous techniques for machine learning, include artificial neural networks, decision trees, and support vector machines (Witten et al., 2005; Saker et al., 2019) are commonly used in the area. An almost infinite several enterprises, organizations, and sectors, such as social networking, government agencies, e-commerce, telecommunications, banking and financial services, healthcare, sales and marketing, and numerous others, can benefit from accurate predictions because they offer insight into the unknown.

C. Analysis of Medical Videos and Images

MVIA stands for Medical Video and Image Analysis. In the healthcare sector, recognizing images remains a challenging task, that in turn lowers the frequency of detection of illnesses and the quality of life (Soguero-Ruiz et al., 2018; Shickel, et al., 2018 & Yin et al., 2017). Artificial intelligence (AI) and machine learning methods can be applied in the healthcare sector to recognize photos and videos produced by machines (Bruzzone & Marconcini, 2010; Fadlullah et al., 2017; Hurlburt, 2017). Machines have become equipped to make judgments and evaluate them similarly to humans. Analysis of medical video and images is completed utilizing the Multiple Instance Learning, or MIL. Multiple Occurrence Learning, or MIL, is used in the evaluation of medical video and image data. Prior to the usage of learning with multiple instances (MIL), supervised learning was employed, whereby every study was assigned a single class label. Currently, in learning with multiple instances (MIL), the observations are gathered, the situations are categorized, groups are created, and class labels are given to these groupings. Utilizing the MIL in the MVIA involves three key tasks. (a) Each bag represents a video or a picture (b) Each instance represents a specific point inside the picture or video, aiding in the extraction of a feature vector (c) The category labels serve as a general diagnostic for the pictures and videos. As a result, MIL on its own recognizes the focus similarities in the images and videos and provides a diagnostic for each one. The feature vectors that are individually retrieved train the classifier that just uses universal identifiers for the control, preventing the patterns in pictures and videos from being mixed. The feature vectors that are individually retrieved only use the global tags to train the classifier as a control preventing the patterns in pictures and videos from being mixed.

D. Telecommunication Systems Using Machine Learning Models

The next part will address the models employed, the methods for gathering training data, and the uses of machine learning in networks of communication. It is possible for artificial intelligence to spot trends in unstructured data and learn new information, a process known as machine learning (86). The domains of image identification and natural language processing have benefited greatly from this technique (86). The projections get more precise as the data set grows. Although the field of machine learning is still very young in the area of communication networks, there is little question that subsequently will be important in the years to come. In networks that are IP, mobile, optical, or IOT, there are tens of thousands of network components, and machine learning can handle this enormous volume of data. Communication networks acquire a vast amount of data since there are several distinct device kinds with numerous manufacturers, all of which provide data of various forms and formats. As a result, the data gathered is vast and quite diversified. Time is commonly employed to connect data from many sources, such it is crucial to record the timestamp each time you compute data.

7. Conclusion

In conclusion, the exploration of diverse machine learning methods for facies classification represents a pivotal juncture in the evolution of geosciences and the broader field of data-driven decision-making. This journey into the landscape of machine learning in facies classification has illuminated several key takeaways.

Firstly, it is evident that the integration of machine learning techniques possesses the capacity to revolutionize the way geoscientists analyze and interpret subsurface reservoirs. The speed, accuracy, and capability to handle complex multi-modal data are undeniable advantages that can significantly enhance our understanding of geological formations

However, the path forward is not without its challenges. The need for high-quality, labeled datasets remains a primary hurdle, demanding substantial investments in data acquisition and curation. Furthermore, ensuring the interpretability and trustworthiness of machine learning models is paramount. Developing methods to explain model decisions and mitigate biases in training data are areas of active research that require sustained attention.

As we journey further into this landscape, it is essential for stakeholders in the geosciences to collaborate, share knowledge, and establish best practices. Collaboration between domain experts, data scientists, and industry professionals can accelerate progress and foster innovation.

In essence, exploring the landscape of diverse machine learning methods for facies classification is not a one-time expedition but an ongoing endeavor. As we move forward, addressing the identified challenges and embracing emerging technologies will enable us to unlock new insights into subsurface reservoirs and better equip ourselves to make informed decisions in the ever-evolving field of geosciences. The journey continues, and the potential for discovery remains boundless.

In addition to providing a thorough understanding of how cutting-edge AI techniques are revolutionizing the characterization of subsurface reservoirs, this investigation into "Unveiling Dominant Approaches, Benefits, and Potential Pitfalls in Utilizing Diverse Machine Learning Methods for Facies Classification" highlights the synergy between contemporary machine learning techniques and geoscience. Important points to note are:

Cutting-Edge Approaches: Identifying the newest and best machine learning techniques for classifying facies and offering engineers and geoscientists a road map for utilizing AI in the interpretation of geological data.

Tangible Benefits: Outlining the concrete benefits of using machine learning in this situation, such as improving efficiency and accuracy and optimizing resource allocation, which has an immediate influence on resource management and decision-making.

Navigating Challenges: A detailed analysis of the possible obstacles and traps that will provide practitioners with the knowledge they need to reduce risks and maximize model performance when working with complicated geological data.

Benchmarking and Selection: A comparative study of several machine learning algorithms to help experts and researchers choose the best models for certain geological and reservoir circumstances.

Real-World Success Stories: These showcase real-world case studies and practical applications where machine learning has revolutionized the categorization of facial features, offering guidance and motivation to those who are thinking about using it.

Direction and Guidance: Providing useful suggestions and directions for the successful application of machine learning so that engineers and geoscientists may make decisions tailored to their particular project needs.

Pioneering the Future: Investigating the boundaries of machine learning in reservoir engineering and geoscience, offering creative directions for further study, and incorporating cutting-edge technology to promote the area.

The highlight of this topic summarizes how important it is to use a variety of machine learning techniques to manage resources more effectively, decipher complicated geological processes, and spur innovation in the energy and natural resources industries.

References

[1] Wang, Ai, X. and B. Sun, (2019). Automatic identification of sedimentary facies based on a support vector machine in the Aryskum Graben, Kazakhstan. Appl. Sci. 9, 4489 (2019).

[2] R. A, Akkurt, T. Tim, D. P., Conroy, P., Andrea, L, Jacob, and S. Paul, (2018). "Accelerating and enhancing petrophysical analysis with machine learning: a case study of an automated system for well log outlier detection and reconstruction. SPWLA 59th Annual Logging Symposium, One Petro (2018).

[3] A. Al-Anazi, and I.D. Gates (2009a). On the capability of support vector machines to classify lithology. Accepted, November, 2009. Natural Resources Research.

[4] A. Al-Anazi, and I.D. Gates (2009b). Support vector regression for permeability prediction in a heterogeneous reservoir: a comparative study. Accepted, October, 2009. SPE Reservoir Evaluation and Engineering: Formation Evaluation.

[5] A. Al-Anazi, and I.D. Gates (2009c). Support vector regression for porosity prediction in a heterogeneous reservoir: a comparative study. Submitted to Computers and Geosciences.

[6] K. Alexander, S. Pavel, Z. Vladimir, I. Sergey, K. Alexey, K. Maxim, and I. Semen (2019). Field-Scale Automatic Facies Classification Using Machine Learning Algorithms. Paper presented at the SPE Russian Petroleum Technology Conference, Moscow, Russia, October 2019. Paper Number: SPE-196860-MS

[7] J.K. Ali (1994). Neural networks: A new tool for the petroleum industry? in: Society of Petroleum 1167 Engineers - European Petroleum Computer Conference 1994, EPCC 1994. Society of Petroleum 1168 Engineers, pp. 233–242.

[8] S. Anirbid, Y. Kriti, R. Kamakshi, B. Namrata, and O. Hemangi (2021). Application of machine learning and artificial intelligence in oil and gas industry, Petroleum Research, Volume 6, Issue 4, 2021, Pages 379-391.

[9] R. Ashena and G. Thonhauser (2015). Application of artificial neural networks in geoscience and petroleum industry. Artificial Intelligent Approaches in Petroleum Geosciences, Springer (2015), pp. 127-166.

[10] C. M. Bishop (2006). Pattern Recognition and Machine Learning. Springer-Verlag, New York.

[11] S. B. Bakdi, K. Nitish, R.M. Shashipal, and C. Balaji, (2020). Automated well correlation using machine learning and facial recognition techniques. Abu Dhabi International Petroleum Exhibition & Conference, One Petro (2020).

[12] J. L. Baldwin, R.M. Bateman, and C.L. Wheatley (1990). Application of a neural network to the problem of mineral identification from well logs. The Log Analyst, 31(05).

[13] S. Basu, K. Johansen and S. Sergey (2020). "Petrophysical Workflows vs Machine Learning – A Comparative Case Study in the Dakota Group, Williston Basin." Paper presented at the SPWLA 61st Annual Logging Symposium, Virtual Online Webinar, June 2020.

[14] P. Bestagini, L. Vincenzo and T. Stefano (2017). "A machine learning approach to facies classification using well logs", SEG Technical Program Expanded Abstracts 2017.

[15] V. Bolandi, A. Kadkhodaie and R. Farzi, (2017). Analyzing organic richness of source rocks from well log data by using SVM and ANN classifiers: a case study from the Kazhdumi formation, the Persian Gulf basin, offshore Iran. J. Pet. Sci. Eng. 151, 224–234 (2017).

[16] M.S.M. Bolton (2020). Machine learning classifiers for attributing tephra to source volcanoes: an evaluation of methods for Alaska tephras. J. Quat. Sci. 35, 81–92 (2020).

[17] S.B. Brazell, B. Alex, A. Michael and B. Darrin (2019): A machine-learning-based approach to assistive well-log correlation. Petrophysics-The SPWLA Journal of Formation Evaluation and Reservoir Description, 60 (2019), pp. 469-479 04.

[18] L. Breiman (1996). "Stacked regressions". Machine Learning. 24: 49–64. doi:10.1007/BF00117832.

[19] L. Breiman (1996). Bagging Predictors, Machine Learning, 24(2), pp.123-140, 1996.

[20] L. Breiman, (2001). Machine Learning, 45(1), 5 32. Kluwer Academic Publishers. Manufactured in The Netherlands

[21] G. Brown (2010). "Ensemble learning in Encyclopedia of Machine Learning, pp.312–320, Springer, New York, NY, USA, 2010.

[22] L. Bruzzone and M. Marconcini (2010). Domain Adaptation Problems: A DASVM Classification Technique and a Circular Validation Strategy. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(5), 770-787

[23] L. Buşoniu, R. Babuška and B. De Schutter (2010). Multi-agent Reinforcement Learning: An Overview. In: Srinivasan, D., Jain, L.C. (eds) Innovations in Multi-Agent Systems and Applications - 1. Studies in Computational Intelligence, vol 310. Springer, Berlin, Heidelberg.

[24] L. Cao (2017). Data science: a comprehensive overview. ACM Computer Survey (CSUR). 2017;50(3):43.

[25] R.D.M. Cappelli and D. Maltoni (2002). "A multi-classifier approach to finger print classification," Pattern Analysis and Applications, vol.5, no.2, pp. 136–144,2002.

[26] A. Chamkalani, S. Zendehboudi, R. Chamkalani, A. Lohi, A. Elkamel and I. Chatzis (2013). Utilization of support vector machine to calculate gas compressibility factor. Fluid Phase Equilib 358:189–202.

[27] C. Chen, G. Gao, P. Honorio, E. Gelderblom, T. J Jimenez (2014). Integration of principal-component-analysis and streamline information for the history matching of channelized reservoirs SPE Annual Technical Conference and Exhibition, Society of Petroleum Engineers (2014).

[28] C. Choisy, and A. Belaid (2001). Handwriting recognition using local methods for normalization and global methods for recognition. Proceedings of Sixth Int. Conference on Document Analysis and Recognition, pp. 23–27.

[29] T. A. Cross, and P.W. Homewood (1997). "Amanz Gressly's role in founding modern stratigraphy" Geological Society of American Bulletin. Geological Society of America. 109 (12): 1617-1630. Doi; 10.1130/0016-7606.

[30] R. O. Duda, P.E. Hart and D.G. Stork (2001). Pattern Classification. Wiley, New Jersey.

[31] P. Dayan (2009). Unsupervised learning. In The Elements of Statistical Learning; Springer: New York, NY, USA, 2009; pp. 1–7.

[32] P. Delfiner, O. Peyret, and S. Oberto (1987). "Automatic Determination of Lithology from Well Logs." SPE Form Eval 2 (1987): 303–310.

[33] X. Dong, X., GuoJ, and G. Pi. Höök (2015). Sustainability assessment of the natural gas industry in China using principal component analysis Sustainability, 7 (5) (2015), pp. 6102-6118.

[34] Z. M. Fadlullah, F. Tang, B. Mao, N. Kato, O. Akashi,T. Inoue, and K. Mizutani (2017). State-of-the-Art Deep Learning: Evolving Machine Intelligence Toward Tomorrow's Intelligent Network Traffic Control Systems. IEEE Communications Surveys & Tutorials, 19(4), 2432-2455.

[35] M. Fernandez-Delgado, E. Cernadas, S. Barro and D. Amorim (2014). "Do we need hundreds of classifiers to solve real world classification problems? "The Journal of Machine Learning Research, vol.15, no.1, pp. 3133–3181, 2014.

[36] G. Giacinto and F. Roli (2001). "Design of effective neural network ensembles for image classification purposes, "Image and Vision Computing, vol. 19, no. 9-10, pp. 699–707, 2001.
[37] D. Gao, J. Zhou, and L. Xin (2001). SVM-based detection of moving vehicles for automatic traffic monitoring. IEEE Intelligent Transportation System, 745–749.
[38] A. Géron, (2017). Hands-on Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems; O'Reilly Media: Sebastopol, CA, USA, 2017.
[39] A. Globerson and N. Tishby (2003). Sufficient Dimensionality Reduction. J. Mach. Learn. Res. 2003, 3, 1307–1331.
[40] C. M. Gifford and A. Agah (2010). Collaborative multi-agent rock facies classification from wireline well log data. Engineering Applications of Artificial Intelligence, vol. 23, pp. 1158-1172, 2010.
[41] I. Goodfellow, Y. Bengio and A. Courville (2016). Deep Learning, MIT.
[42] K-S. Goh and K-T.C. Chang (2001). "SVM binary classifier ensembles for image classification, "in Proceedings of the 10th International Conference on Information and Knowledge Management (ACMCIKM'01), pp. 395–402, November 2001.
[43] I. Guyon and A. Elisseeff (2003). An Introduction to Variable and Feature Selection. J. Mach. Learn. Res. 2003, 3, 1157–1182.
[44] B. Hall (2016). Facies classification using machine learning. The Leading Edge, 35(10), 906-909.
[45] J. Han, J. Pei, and M. Kamber (2011). Data mining: concepts and techniques. Amsterdam: Elsevier; 2011.
[46] M. Han. and B. Liu (2015) "Ensemble of extreme learning machine for remote sensing image classification" Neurocomputing", vol. 149, pp. 65–70, 2015.
[47] M. Hiren, R. Srikanth., S.K. Mandar, and A. Aria (2018). "Machine learning methods in geoscience", SEG Technical Program Expanded Abstracts 2018.
[48] G. Hurlburt (2017). How Much to Trust Artificial Intelligence? IT Professional, 19(4), 7-11.
[49] J. L. Jing and C. L. Jian (2022). "Integrating deep learning and logging data analytics for lithofacies classification and 3D modeling of tight sandstone reservoirs", Geoscience Frontiers, vol. 17, no. 1, 2022.
[50] M. R. Khan, Z. Tariq, and A. Abdulazeez (2018). "Machine Learning Derived Correlation to Determine Water Saturation in Complex Lithologies." Paper presented at the SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition, Dammam, Saudi Arabia, April 2018.
[51] A. Kharbouch, A. Shoeb, and J. Guttag, (2011). An algorithm for seizure onset detection using intracranial EEG. Epilepsy Behave. 22 (Suppl. 1), S29–S35.
[52] K. Kim, K. Jung, S. Park and H.J. Kim (2001). Support vector machine-based text detection in digital video. Pattern Recognition 34, 527–529.
[53] O. Koteluk, A. Wartecki, S. Mazurek, I. Kołodziejczak, and A. Mackiewicz (2021). How Do Machines Learn? Artificial Intelligence as a New Era in Medicine. Journal of Personalized Medicine. 2021; 11(1):32.
[54] T. Kuwatani (2014). Machine-learning techniques for geochemical discrimination of 2011 Tohoku tsunami deposits. Sci. Rep. 4, 7044, 2014.
[55] L.I. Kuncheva, M. Skurichina, and R. P. W Duin, (2002). "An experimental study on diversity for bagging and boosting with linear classifiers," Information Fusion, vol. 3, no. 4, pp. 245–258, 2002.
[56] A. Leila, O. Sid-Ali, D. Nouredine and B. Amar (2013). "Lithofacies prediction from well logs data using different neural network models", Proceedings of the 2nd International Conference on Pattern Recognition Applications and Methods (PRG-2013), 2013.
[57] J. Li, K. Cheng, S. Wang, F. Morstatter, R.P. Trevino, J. Tang, and H. Liu (2017). Feature Selection: A Data Perspective. ACM Computer. Survey. 2017, 50-94.
[58] L. I. Kuncheva, Skurichina, M., and Duin, R. P. W., (2002). "An experimental study on diversity for bagging and boosting with linear classifiers," Information Fusion, vol. 3, no. 4, pp. 245–258, 2002.
[59] Z. Li, Z. Weida and J. Licheng (2000). Radar target recognition based on support vector machine. Proceedings of 5th Int. Conference on Signal processing, vol. 3, pp. 1453–1456.
[60] X. Li, C.W. Chan and H. H. Nguyen (2013). Application of the Neural Decision Tree approach for prediction of petroleum production. J Pet Sci Eng 104:11–16.
[61] J. Lu, K. Plataniotis and A. Ventesanopoulos (2001). Face recognition using feature optimization and v-support vector machine. IEEE Neural Networks for Signal Processing XI, 373–382.
[62] Lui, Hui-Hai Liu, Z. Jilin, L. Feng, C. Temizel, A. B. Mustafa and M. Rabah (2021)."Incorporation of physics into machine learning for production prediction from unconventional reservoirs: a brief review of the gray-box approach" SPE Reservoir Eval. Eng., 24 (2021), pp. 847-858.
[63] C. Ma, M. Randolph, and J. Drish (2001). A support vector machines-based rejection technique for speech recognition. Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing 1, 381–384.
[64] M.S. Mahdavinejad, M. Rezvan, M. Barekatain, P. Adibi, P. Barnaghi, and A. P. Sheth (2018). Machine learning for internet of things data analysis: a survey. Digit Commun Netw. 2018;4(3):161–75.
[65] S. Ma and A. Ma (2015). Core analysis with emphasis on carbonate rocks-quality assurance and control for accuracy and representativeness. SEG Interpretation Journal, 3(1), 2015.
[66] I. Marco, J. F. Ferguson and J. Fred (2021). Improving facies prediction by combining supervised and unsupervised learning methods (2021). Journal of Petroleum Science and Engineering (Elsevier), vol. 200, pp 108300.
[67] R. Mehran and A. R. Mohammad (2022). Reservoir facies classification based on random forest and geostatistics methods in an offshore oilfield. Journal of Applied Geophysics, vol. 201, pp. 104640-104640.
[68] A. Merentitis, C. Debes and R. Heremans (2014), "Ensemble learning in hyper spectral image classification: toward selecting a favorable bias-variance tradeoff," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 7, no. 4, pp. 1089–1102, 2014.
[69] S. M. Misra, O. Oghenekaro and P. Mark (2019). Unsupervised outlier detection techniques for well logs and geophysical data. Mach. Learn. Subsurface Charact, 1, 2019.
[70] S. Mohaghegh., R. Arefi, S. Ameri and D. Rose (1994). Design and development of an artificial neural 1375 network for estimation of formation permeability, in: Proceedings - Petroleum Computer 1376 Conference. Publ by Society of Petroleum Engineers (SPE), pp. 147–154. 1377.
[71] J. Q. Ngui, M. Hermana, D. P. Ghosh and W.L. Wan Yusoff (2018). Integrated study of lithofacies identification—a case study in X field, Sabah, Malaysia. Geosciences, 8(2), 75.
[72] M. Ozay and V. F. T. Yarman, (2013). "A New Fuzzy Stacked Generalization Technique and Analysis of its Performance".
[73] M. Pal (2008). "Ensemble of support vector machines for landcover classification," International Journal of Remote Sensing, vol. 29, no. 10, pp. 3043–3049, 2008.
[74] S. P. Parker (1984). McGraw-Hill concise Encyclopedia of science and technology. McGraw-Hill's, 705, ISBN 0-07-045482-5
[75] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer and R. Weiss, V. Dubourg (2011). Scikit-learn: machine learning in python. J Mach Learn Res. 2011; 12:2825–30.
[76] J. R. Quinlan (1993). C4.5: programs for machine learning. Mach Learn. 1993.
[77] R. Ranawana and V. Palade (2006). "Multi-classifier systems: review and a road map for developers," International Journal of Hybrid Intelligent Systems, vol. 3, no. 1, pp. 35–61, 2006.
[78] R. Polikar (2006). "Ensemble based systems in decision making," IEEE Circuits and Systems Magazine, vol. 6, no. 3, pp. 21–45, 2006.
[79] P. Polikar (2007). "Bootstrap-inspired techniques in computation intelligence" IEEE Signal Processing Magazine, vol. 24, no. 4, pp. 59–72, 2007.
[80] H.G. Reading (1996). Sedimentary Environments and Facies. Blackwell Scientific Publications. ISBN 0-632-03627-3.
[81] S. M. Randall, R. Skip, K. Deepak and N. Fernando (2019). "Application of artificial intelligence for depositional facies recognition - Permian basin", Unconventional Resources Technology Conference, Denver, Colorado, USA, 22-24 July 2019.
[82] S. J. Rogers, J. H. Fang, C.L Karr and D. A. Stanley (1992). Determination of lithology from well logs using a neural network. AAPG bulletin, 76(5), 731-739.
[83] L. Rokach (2010). "Ensemble-based classifiers". Artificial Intelligence Review. 33 (1–2): 1–39.
[84] P. Saikia, D. Nankani, and R. D. Baruah (2019) "Reservoir Facies Classification using Convolutional Neural Networks," 2019 IEEE Recent Advances in Geoscience and Remote Sensing: Technologies, Standards and Applications (TENGARSS), Kochi, India, 2019, pp. 34-38.

[85] R. Salman, A. Alzaatreh, H. Sulieman and S. Faisal (2021). A Bootstrap Framework for Aggregating within and between Feature Selection Methods. Entropy (Basel, Switzerland), 23(2), 200.

[86] A. Samat, P. Du, S. Liu, J. Li, and L. Cheng (2014). "E2LMs: Ensemble extreme learning machines for hyper spectral image classification," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 7, no. 4, pp. 1060–1069, 2014.

[87] I. H. Sarker, P. Watters and A.S.M. Kayes (2019). Effectiveness analysis of machine learning classification models for predicting personalized context-aware smartphone usage. J Big Data. 2019;6(1):1–28.

[88] I. H. Sarker, C. Alan, H. Jun, A. I .Khan, Y.B.Abu shark and S. Khaled (2019). Behaved: A behavioral decision tree learning to build user-centric context-aware predictive model. Mob Netw Appl. 2019; 1–11.

[89] I.H. Sarker, H. Alqahtani, F. Alsolami. A. Khan, Y.B Abushark, and M. K. Siddiqui (2020). Context pre-modeling: an empirical analysis for classification-based user-centric context-aware predictive modeling. J Big Data. 2020;7(1):1–23.

[90] I. H. Sarker (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. SN COMPUT. SCI. 2, 160, 2021.

[91] C. Satinder and K. J. Marfurt (2018). "Seismic facies characterization using some unsupervised machine learning methods", SEG Technical Program Expanded Abstracts 2018.

[92] E, S. Satu, Sara and V. R. Sanchez (2020). Forecasting client retention — A machine-learning approach, Journal of Retailing and Consumer Services, vol. 52, 2020, 101918.

[93] B. Seth, B. Alex, A. Michael and B. Darrin (2019). "A machine-learning-based approach to assistive well-log correlation", Petrophysics, vol. 60, no. 4, pp. 469 - 479, 2019.

[94] B. Shickel, P.J. Tighe, A. Bihorac and P. Rashidi (2018). Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis. IEEE Journal of Biomedical and Health Informatics, 22(5), 1589-1604.

[95] P. Smyth and D. Wolpert (1999). "Linearly Combining Density Estimators via Stacking" (PDF). Machine Learning. 36 (1): 59–83.

[96] S. M. Stanley (1999). Earth System History. New York: W.H. Freeman and Company. p. 134. ISBN 0-7167-2882-6.

[97] C. Soguero-Ruiz, K. Hindberg, J. L. Rojo-Alvarez, S. O Skrovseth, F. Godtliebsen, K. Mortensen, and R. Jenssen (2016). Support Vector Feature Selection for Early Detection of Anastomosis Leakage from Bag-of-Words in Electronic Health Records. IEEE Journal of Biomedical and Health Informatics, 20(5), 1404-1415.

[98] C. Sachin and G. P. Karmakar (2020). Artificial intelligence techniques and their application in oil and gas industry. Journal of Artificial Intelligence Review Springer Nature B.V. 2020.

[99] L. Tahiru, E. Maju-Oyovwikowhe and S. Otobo (2021). Application of Machine Learning to Facies Classification from Well Logs.

[100] Tan and Gilbert, (2003): Ensemble machine learning on gene expression data for cancer classification. Appl. Bioinf., 2 (3 Suppl. l) (2003), pp. S75-S83

[101] W. Thilo, R. L. Indranil Pan and H.F. Gawthorpe (2018). "Seismic facies analysis using machine learning", Geophysics, vol. 83, no. 5, 2018.

[102] V. T. Tran, H.H. Ngo, K.S. Hoang, N. Tran and J.J. Lambiase (2020). "Depositional facies prediction using artificial intelligence to improve reservoir characterization in a mature field of Nam Con Son basin, offshore Vietnam", Offshore Technology Conference Asia, Kuala Lumpur, Malaysia, 2-6 November 2020.

[103] T. Van Gestel, J. Suykens, D. Baestaens, A. Lambrechts, G. Lanckriet, B. Vandaele, B. De Moor and J. Vandewalle (2001). Financial time series prediction using least squares support vector machines within the evidence framework. IEEE Transactions. On Neural Networks 12 (4), 809–821.

[104] P. Vladimir and E. Chris (2022). "Unsupervised seismic facies classification using deep convolutional autoencoder", Geophysics, vol. 87, no. 4, 2022.

[105] Yu. Wang, V. T. Igor, A. H. Mark, A.F. Eibe Frank, F. X. Klaus, Mayer, and W. M. Hans (2005). Gene selection from microarray data for cancer classification—a machine learning approach, Computational Biology and Chemistry, vol. 29, no. 1, pp. 37-46, 2005.

[106] I. H. Witten and E. Frank (2005). Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann; 2005.

[107] D. H. Wolpert (1992). "Stacked Generalization". Neural Networks. 5 (2): 241–259.

[108] D. H. Wolpert, MacReady and G. William (1999). "An Efficient Method to Estimate Bagging's Generalization Error" (PDF). Machine Learning. 35 (1): 41–55.

[109] M. Wo´zniak, N. Gra˜ na, and E. Corchado (2014). "A survey of multiple classifier systems as hybrid systems," Information Fusion, vol. 16, no. 1, pp. 3–17, 2014.

[110] T. Wrona, L. Pan, R.L. Gawthorpe and H. Fossen (2018). Seismic facies analysis using machine learning. Geophysics 83, 83–95, 2018.

[111] Y. Xie, C. Zhu, W. Zhou, Z. Li, X. Liu and M. Tu (2018). Evaluation of machine learning methods for formation lithology identification: a comparison of tuning processes and model performances. J. Petrol. Sci. Eng., 160, pp. 182-193, 2018.

[112] H. Yin and N. K. Jha (2017). A Health Decision Support System for Disease Diagnosis Based on Wearable Medical Sensors and Machine Learning Ensembles. IEEE Transactions on Multi Scale Computing Systems, 3(4), 228-241.

[113] C. Zhang and Y. Ma (2012). Ensemble machine learning, Springer, NY, USA, 2012.

[114] Z. H. Zhou (2012). Ensemble Methods: Foundations and Algorithms, Chapman &Hall/CRC, 2012.

[115] Zhou and Belkin, (2014). Semi-Supervised Learning, Academic Press Library in Signal Processing, Elsevier, Amsterdam, pp. 1239-1269, 2014.