

Research
Unconventional and Intelligent Oil & Gas Engineering—Perspective

Intelligent Petroleum Engineering

Mohammad Ali Mirza, Mahtab Ghoroori*, Zhangxin Chen*

Department of Chemical and Petroleum Engineering, University of Calgary, Calgary, AB T2N 1N4, Canada



ARTICLE INFO

Article history:

Received 9 December 2021

Revised 1 May 2022

Accepted 17 June 2022

Available online 19 July 2022

Keywords:

Artificial intelligence

Machine learning

Intelligent reservoir engineering

Text mining

Intelligent geoscience

Intelligent drilling engineering

ABSTRACT

Data-driven approaches and artificial intelligence (AI) algorithms are promising enough to be relied on even more than physics-based methods; their main feed is data which is the fundamental element of each phenomenon. These algorithms learn from data and unveil unseen patterns out of it. The petroleum industry as a realm where huge volumes of data are generated every second is of great interest to this new technology. As the oil and gas industry is in the transition phase to oilfield digitization, there has been an increased drive to integrate data-driven modeling and machine learning (ML) algorithms in different petroleum engineering challenges. ML has been widely used in different areas of the industry. Many extensive studies have been devoted to exploring AI applicability in various disciplines of this industry; however, lack of two main features is noticeable. Most of the research is either not practical enough to be applicable in real-field challenges or limited to a specific problem and not generalizable. Attention must be given to data itself and the way it is classified and stored. Although there are sheer volumes of data coming from different disciplines, they reside in departmental silos and are not accessible by consumers. In order to derive as much insight as possible out of data, the data needs to be stored in a centralized repository from where the data can be readily consumed by different applications.

© 2022 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Artificial intelligence (AI) seeks to mimic human decision making. The subset of AI techniques known as machine learning (ML) enables computers to learn how to act outside the confines of their programmed behavior, through the use of external data. ML has revolutionized myriad industries and fields of study, with incredibly wide-ranging applications from stock market analysis to self-driving cars. With the advent of Internet of Things (IoT) devices and big data (where a high volume of data is generated at high velocity and with many different varieties), ML is one of the most important technologies to ensure that actionable insights can be gleaned from big data.

In the oil and gas industry, model types are divided into three main categories: physical, mathematical, and empirical models [1]. A physical model is a scaled-down or scaled-up version of an object that is developed to simplify the understanding of how a physical object or scenario looks or operates. These models have the disadvantages of being costly and time-consuming to develop,

and may not be sufficiently accurate in some cases. Empirical models are established based on experiments; they are subject to a variety of errors, such as human and measurement errors, and are not generalizable. Mathematical models encode physical laws to simulate the underlying physics; however, they require many assumptions and simplifications [1]. To deal with the challenges in these three model types, derive insights, and make intelligent decisions in a timely manner, a more promising technique is required. This is where ML can be applied, due to its ability to capture and act upon insights from vast datasets that could never be handled through purely programmatic rules, due to the complexity of the relationships between data and the insights gleaned from the data.

The oil and gas industry is rapidly transitioning to oil-field digitization, and there has been an increased drive to apply data-driven modeling and ML algorithms to various petroleum engineering challenges. Data-driven modeling uses mathematical equations derived from data analysis, as opposed to knowledge-driven modeling, in which logic is the main tool to represent a theory [2,3]. While there may be data-driven algorithms that do not learn from data (and thus cannot be called ML), ML is a subset of data-driven approaches that demonstrate a form of AI. Fig. 1 summarizes different types of ML algorithms.

* Corresponding authors.

E-mail addresses: mahtab.ghoroori@ucalgary.ca (M. Ghoroori), zhachen@ucalgary.ca (Z. Chen).

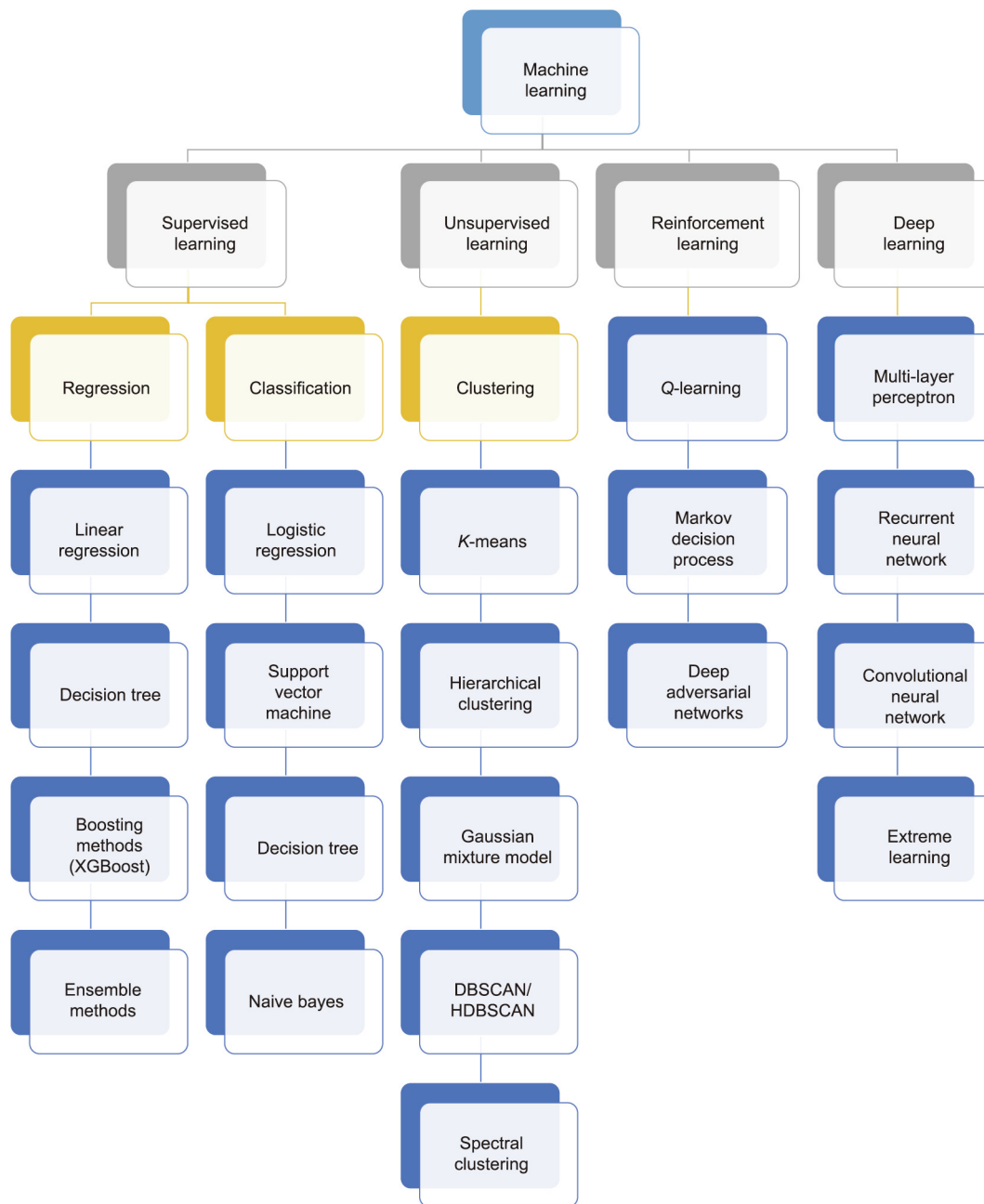


Fig. 1. Different types of ML algorithms. DBSCAN: density-based spatial clustering of applications with noise; HDBSCAN: hierarchical density-based spatial clustering of applications with noise.

ML has been widely used in different areas of the petroleum industry, including geoscience, reservoir engineering, production engineering, and drilling engineering. The next four sections present a critical review and perspective of the application of ML in each of these areas.

1.1. Intelligent geoscience

Geoscience has utilized algorithms such as decision trees, Markov chains, and *K*-means clustering as early as the 1960s. Markov chains have been utilized in sedimentology [4], hydrology [5], and well-log analysis [6]. Preston and Henderson [7] used *K*-means clustering to interpret the periodicity of sediment deposits. Early application of decision trees can be found in economic geology and perspective mapping [8,9]. Due to a variety of factors,

including a lack of computational power and immaturity of the field, ML adoption did not perfectly satisfy initial expectations; hence, little development of AI occurred in the 1970s. Zhao and Mendel [10] employed recurrent neural networks (NNs) to perform seismic deconvolution in the 1980s, which can be considered a resurgence of interest in AI. A shift from knowledge-driven to data-driven ML occurred in the 1990s, when the first review of NNs in geophysics was published by McCormack [11]. McCormack’s review explored pattern recognition and presented a summary of NN applications over the previous 30 years, along with applied samples of seismic trace editing and automated well-log analysis. Deep learning (DL) and, more specifically, convolutional neural networks (CNNs) were revitalized in the 2010s, when Waldeland and Solberg [12] applied a small CNN to seismic data for salt recognition. Fault interpretation [13–15], horizon

picking [16], and facies classification [17,18] are other applications of CNNs in geoscience. Mosser et al. [19] applied a generative adversarial network (GAN)—that is, an ML model in which two NNs work together competitively to make more accurate predictions—early on in geoscience to perform pore scale modeling of three-dimensional (3D) porous media. GANs have also been utilized in seismogram generation and geostatistical inversion [20].

Seismicity is another important field in geoscience in which ML has become widely used. Mousavi et al. [21] used ML algorithms to discriminate deep micro-seismic events from shallow ones based on features of waveforms recorded on surface receivers. He et al. [22] used an ML algorithm to improve the risk management of induced seismic events. The proposed model was a set of simple closed-form expressions, with the advantages of high transparency and fast execution speed, providing the operator with the greatest chance of success. Industrial activities such as mining, oil- and gas-field depletion, wastewater injection, and geothermal operations can induce seismicity [23,24]. In western Canada, the seismicity induced from hydraulic fracturing (HF) has galvanized public and academic attention [25]. Investigating the correlations between induced seismicity and HF has been exceedingly challenging for a long time, due to the complexity introduced by strongly coupled geomechanical, geophysical, and geological behaviors. Thus, there is plenty of room for exploring ML applications in seismicity.

1.2. Intelligent reservoir engineering

ML algorithms have become popular in various areas of reservoir engineering, particularly in reservoir characterization and in pressure, volume, and temperature (PVT) computations. A two-layer NN was developed by Gharbi and Elsharkawy [26] to estimate bubble point pressure and a formation volume factor for oil reservoirs. In another study, a radial basis function and multi-layer perceptron NN were employed to estimate a formation volume factor, isothermal compressibility, and brine salinity [27]. Wang et al. [28] used artificial neural networks (ANNs) in a compositional reservoir simulation for phase equilibrium calculations, including phase stability tests and phase splitting calculations. A combination of two approaches—namely, a support vector machine (SVM) and fuzzy logic—was utilized to predict permeability and porosity using real-life well logs as an input feed [29]. Patel and Chatterjee [30] utilized classification algorithms to carry out quick and accurate rock typing (i.e., classifying reservoir rock into different categories based on similarities). In the presence of random noise, the performance of an ANN with a single hidden layer was explored by An [31] to establish a model to predict the thickness of a low-velocity layer. The proposed approach was also applied on an oil field in northern Alberta, Canada, to construct a distribution map of porosity-net pay thickness, based on which four wells were drilled and the field production increased by almost 20% [32]. Jamialahmadi and Javadpour [33] utilized a radial basis function NN using depth measurements and the porosity of core data as inputs to estimate the permeability of an entire oil field in southern Iraq. An ensemble ML model (i.e., a random forest algorithm) was developed by Wang et al. [34] to predict time-lapse saturation profiles at well locations using actual production and injection data from a structurally complicated and highly faulted offshore oil field as the major inputs. A new framework for the prediction of multiple reservoir parameters (i.e., porosity, saturation, lithofacies, and shale content) was developed by introducing an extreme learning machine (ELM), which is one of the most advanced ML techniques [35]. In comparison to the classic single-layer feed-forward NN approach, the proposed method requires fewer computing resources and less training time without sacrificing accuracy.

1.3. Intelligent production engineering

Production prediction/optimization and HF are two other fields in the energy industry in which ML has grown popular. Many parameters must be taken into account for production prediction and optimization, including the recovery process, proppant type, well spacing, treatment rate, and number of fracturing stages. Although the optimization of operational parameters can save millions of dollars and significantly enhance unconventional reservoir production, traditional reservoir simulations are computationally expensive, which adds up when taking different variations of reservoir characteristics into account [36,37]. Hence, production prediction and optimization are good candidates for AI applications, as shown by the recent development and analyses of ML algorithms for various recovery processes, such as water and chemical floods and steam injection [38–40]. Dang et al. [41] utilized an NN for the multidimensional interpolation of relative permeability to overcome the impacts of different parameters (i.e., the polymer, surfactant, and salinity) during hybrid recovery processes. Production forecasting for wells in different reservoirs using geological, core, and log data is a widely used ML application in this domain [42,43]. Tadjer et al. [44] utilized DeepAR and Prophet (a time series ML algorithm) as alternatives to decline curve analysis for short-term oil and gas well forecasting. Using an NN to predict bottom hole pressure in vertical wells, which is a crucial parameter in the design of production facilities, is another application of ML in this area [45]. A long short-term memory (LSTM) model along with a feature-selection method was applied to predict the daily production rates of shale gas wells in the Duvernay Formation in Canada [46]. Horizontal well placement optimization was investigated by Popa and Connel [47] via stratigraphic performance estimation using a combination of fuzzy logic and NN.

In the last two decades, the growing number of HF jobs has resulted in a substantial amount of measured data that can be used to construct ML prediction models. A study was conducted by Mohaghegh [48] to map a natural fracture network in Utica shale using fuzzy logic cluster analysis. He et al. [49] developed a model to optimize HF design in shale gas reservoirs using AI and fuzzy logic analysis. A novel model was developed using an SVM to determine the hydraulic apertures of rough rocks [50]. Yang et al. [51] established a data analytics approach combining design parameters derived from acoustic wireline logs and post-fracturing analysis to optimize fracturing treatment design. The obtained fracturing optimization algorithm was validated using production logging tool data and deep shear-wave imaging along horizontal wells in the Marcellus shale reservoirs. An integrated approach combining ML, reservoir simulation, and HF was presented by Wang and Sun [52] to optimize well spacing in Permian shales, considering a typical well for each representative region in this large area. A reinforcement learning algorithm was applied by Bangi and Kwon [53] to achieve a uniform proppant concentration along fractures in order to improve HF productivity; the researchers coupled dimensionality reduction with transfer learning to speed up the learning process. Duplyakov et al. [54] presented a model based on a combination of boosting algorithms and ridge regression to predict the cumulative oil production of a well completed with multistage fractures. A case study was performed on 74 hydraulically fractured wells in the Montney Formation in Alberta, Canada, to predict cumulative production profiles over a five-year period using well spacing, rock mechanical properties, and completion parameters as input features [55]. A proxy model was developed to predict cumulative gas production for shale reservoirs using a deep NN based on production, completion, and HF data as input features; this model was validated using field data for 1239 horizontal wells in the Montney Formation [56].

1.4. Intelligent drilling engineering

Considering that huge volumes of real-time data are being produced daily during drilling operations, drilling engineering has also benefited greatly from the application of ML. Due to the volatility of oil prices in recent years, methods to ensure good economics in a variety of price scenarios have been sought. In order to achieve this, ML has become increasingly common to alleviate drilling challenges in real time. Drilling operation optimization and stuck pipe prediction are two of the most critical areas in drilling engineering and have been frequently investigated using AI applications. Mohaghegh [57] utilized an ANN for the real-time identification of drilling anomalies and their related nonproductive time (NPT). An ML model was developed by Unrau et al. [58] to determine a real-time alarm threshold in order to detect anomalies in flow rates and mud volume data during drilling operations. This model assists in the early detection of lost circulation and minimizes false alarm creation. Reinforcement learning algorithms were applied by Pollock et al. [59] to refine a pretrained NN based on 14 horizontal wells in the Permian and Appalachia basins. The refined model managed to minimize tortuosity and deviations from planned trajectories with a less than 3% error. Zhao et al. [60] applied ML algorithms to derive a trend of different drilling parameters in order to identify anomalous incidents and propose remedial actions accordingly. An attempt was made to apply ML algorithms to the optimization of a rate of penetration (ROP) using drilling features such as weights on bit, flow rate, and rotations per minute [61]. Goebel et al. [62] developed an ML model to predict future stuck pipes based on the monitoring and investigation of various parameters including ROP, pipe rotation, inclination angle, and flow rate. A year later, real-time risk prediction during drilling was presented by Dursun et al. [63]. ML algorithms were coupled with data mining and natural language processing (NLP) techniques to investigate daily drilling reports (DDRs) for two onshore fields in the Middle East in an exceptionally short time, in order to categorize productive and NPT and discover critical contributing factors of NPT [64].

2. Challenges and opportunities

ML algorithms can be very effectively applied to address three main types of problems: building surrogate models for understood problems to reduce computational costs; building ML models for problems that require human intervention and knowledge for analysis; and building ML models for complex problems that were previously impractical to be addressed. ML yields the fastest success in realms in which the environment is straightforward, data is easily available, and decisions are not expensive. Although most cases of ML use in the petroleum industry do not meet any of these criteria, as the environments are usually heterogeneous, decisions are expensive (e.g., drilling a well), and data is sporadic, the investment in longer term gains through the effective application of ML can provide a great deal of value, although thoughtful design and a high degree of collaboration with domain experts is required [65].

Applying ML algorithms to petroleum engineering problems requires a variety of challenges to be overcome. One challenge is that the data often comes in a high volume (i.e., large amounts of data), with a wide range of variety (i.e., many different data formats) and veracity (i.e., data inconsistency and inaccuracy), and at a high velocity (i.e., a high rate of data influx). Massive amounts and varieties of data are being produced daily from downhole and surface sensors installed on operational equipment in the petroleum industry. The industry utilizes structured and unstructured data to keep track of production, safety, and maintenance. Acquiring accurate data in the petroleum industry is usually difficult or

impossible, and can be expensive. As a result, obtaining sufficient quantities of high-quality data for training and verification is a prevalent challenge in the petroleum industry, which causes uncertainties and noise in training data. In turn, such issues compromise the generalizability and accuracy of ML models. In addition, raw data is often not ready for ML algorithms and needs to be preprocessed and cleaned. Subsurface uncertainties and data-processing time delays are also important considerations. Moreover, such data usually resides in departmental silos, and the corresponding models are either unavailable or not open to others due to confidentiality concerns and competitive edges—a problem that is particularly prevalent in academic environments [65]. Furthermore, model explainability is important for geoscience, since it can be just as important to know the reason for a result as it is to know the result itself.

Perhaps due to the challenges mentioned above, ML adoption in geoscience is not moving as quickly as in many other fields. Although ML is a promising technique for using big data to discover input–output relationships and derive insights, ML performance can be affected by the high dimensionality of the data. This may lead to misleading correlations and impractical and unreliable clustering. It is noteworthy that data is usually ambiguous in its initial state; thus, different preprocessing techniques are required to identify salient features and make the ML model capable of learning a system's behavior. There is a risk that missing data and a lack of system stability will introduce biases into ML models, making it problematic to extract beneficial knowledge from data [66]. Moreover, the considerations and challenges of utilizing data from diverse sources should be taken into account. Privacy, security, and ethics related to data are also very important aspects to consider. Hybrid modeling, which integrates ML algorithms with physics-based methods, can be considered as a way to mitigate the abovementioned problems. Furthermore, transfer learning—in which a pretrained model is used as a starting point, and then a model is trained on top of it by considering one's own training data—is a relatively recent ML technique that can potentially be beneficial in geoscience contexts.

3. Perspectives

The potential of ML has not been fully used in two areas of the petroleum industry—namely, reservoir simulation and text mining. Reservoir simulation involves differential equations (DEs) that adequately illustrate physical property changes over time and space and are thus useful for describing physical phenomena in nature. There are many problems in science and engineering that require solving complicated DEs. However, DEs are remarkably difficult to solve, and their associated simulations are extremely complex and computationally intensive. This level of complexity requires the use of giant computers to perform simulations and justifies the interest in AI among researchers in this area. Utilizing DL, which involves NNs with more than one hidden layer, is a promising technique that will speed up solving DEs and save scientists and engineers a great deal of time and effort. Caltech researchers have introduced a new DL technique for solving DEs that is more accurate, generalizable, and 1000 times faster than traditional DL algorithms [67]. This new adaptation is based on defining the input and output in a Fourier space, as opposed to a Euclidean space in traditional DL. This development will not only lessen the dependency on supercomputers but also raise the computational capacity to efficiently model more intricate problems.

The petroleum industry is just beginning to harness the power of ML for smart reporting and extracting information from text documents. Daily drilling and completion reports are two of the main text-based documents in the industry that contain important

text, as well as a variety of other types of data such as depths, casing sizes, hole sizes, and perforation depths. NLP and DL algorithms can be used to develop models for automated quality control of operations and performance improvement, providing approaches that are far more efficient than the traditional approach of relying on the knowledge of subject matter experts [68]. Several studies have investigated text processing in the petroleum industry, with a focus on topics including the text mining of operational data for risk management and issue prediction [69], producing metrics and pattern recognition based on contextual analysis of reports [70,71], and reports classification [71]. Although the literature contains such studies on using text mining techniques to mitigate text-based challenges in the industry, there is still great potential for ML in this area, and it must be further explored.

4. Concluding remarks

Data-driven approaches and AI algorithms hold enough promise that they may someday be relied upon even more than physics-based methods. Their main feed is data, which is the fundamental element of each scenario. These algorithms learn from data and reveal unseen patterns. Within the petroleum industry, there is great interest in using this technology to gain insight from the huge volumes of data that are generated every second. Many studies explore AI applicability in various subdisciplines of this industry; however, there is a noticeable lack of two main features; that is, most of the research on this topic is either not practical enough to be applicable in real-field challenges or limited to a specific problem and not generalizable. Attention must be given to data itself and to how it is classified and stored. Although tremendous volumes of data are produced in different disciplines, such data remains within departmental silos and is not accessible by others. To derive as much insight as possible from data, the data must be stored in a centralized repository from which it can be readily consumed for different applications. Between data acquisition and the application of AI and ML techniques, data must be processed in order to effectively extract features and ensure that the data can effectively support the algorithms. Although AI and ML techniques are increasingly important within the petroleum and reservoir engineering domain, they are only part of a holistic system. In order to ensure that this system can deliver value, careful consideration is required to apply algorithms to this challenging domain, and the right type, quality and volume of data must be available and effectively processed to achieve the desired outcomes. Thus, although AI is a critical tool to efficiently manage the world's underground resources, data is the key to fully exploit the possibilities.

Compliance with ethics guidelines

Mohammad Ali Mirza, Mahtab Ghoroori, and Zhangxin Chen declare that they have no conflict of interest or financial conflicts to disclose.

References

- [1] Noshi CI, Schubert JJ. The role of machine learning in drilling operations: a review. In: Proceedings of the SPE/AAPG Eastern Regional Meeting; 2018 Oct 7–11; Pittsburgh, PA, USA. Richardson: OnePetro; 2018.
- [2] Solomatine DP, Ostfeld A. Data-driven modelling: some past experiences and new approaches. *J Hydroinform* 2008;10(1):3–22.
- [3] Dubois D, Hájek P, Prade H. Knowledge-driven versus data-driven logics. *J Logic Lang Inform* 2000;9(1):65–89.
- [4] Schwarzbacher W. The semi-Markov process as a general sedimentation model. In: Merriam DF, editor. Mathematical models of sedimentary processes: an international symposium. Boston: Springer; 1972. p. 247–68.
- [5] Matalas NC. Mathematical assessment of synthetic hydrology. *Water Resour Res* 1967;3(4):937–45.
- [6] Agterberg FP. Markov schemes for multivariate well data. In: Proceedings of the International Symposium on the Application of Computers and Operations Research in the Mineral Industry. Pennsylvania: Pennsylvania State University; 1966.
- [7] Preston FW, Henderson J. Fourier series characterization of cyclic sediments for stratigraphic correlation. In: Merriam DF, editor. Proceedings of the Symposium on Cyclic Sedimentation. Kansas: Geological Survey; 1964. p. 415–25.
- [8] Newendorp PD. Decision analysis for petroleum exploration. Tulsa: PennWell Books; 1976.
- [9] Reddy RKT, Bonham-Carter GF. A decision-tree approach to mineral potential mapping in Snow Lake Area. *Manitoba Can J Rem Sens* 1991;17(2):191–200.
- [10] Zhao X, Mendel JM. Minimum-variance deconvolution using artificial neural networks. In: SEG technical program expanded abstracts 1988. Houston: SEG Library; 1988.
- [11] McCormack MD. Neural computing in geophysics. *Lead Edge* 1991;10(1):11–5.
- [12] Waldebrand AU, Solberg AHSS. Salt classification using deep learning. In: Proceedings of the 79th EAGE conference and exhibition 2017; 2017 Jun 12–15; Paris, France; 2017.
- [13] Araya-Polo M, Dahlke T, Frogner C, Zhang C, Poggio T, Hohl D. Automated fault detection without seismic processing. *Lead Edge* 2017;36(3):208–14.
- [14] Guitton A. 3D convolutional neural networks for fault interpretation. In: Proceedings of the 80th EAGE conference and exhibition 2018; 2018 Jun 11–14; Copenhagen, Denmark; 2018.
- [15] Purves S, Alaei B, Larsen E. Bootstrapping machine-learning based seismic fault interpretation. In: Proceedings of the AAPG Annual Convention and Exhibition; 2018 May 20–23; Salt Lake City, UT, USA; 2018.
- [16] Wu H, Zhang B. Semi-automated seismic horizon interpretation using encoder-decoder convolutional neural network. In: SEG technical program expanded abstracts 2019. Houston: SEG Library; 2019.
- [17] Chevatarese DS, Szwarcman D, Gama e Silva RM, Vital Brazil E. Deep learning applied to seismic facies classification: a methodology for training. In: Proceedings of the European Association of Geoscientists & Engineers, Saint Petersburg 2018; 2018 Apr 9–12; Saint Petersburg; 2018. p. 1–5.
- [18] Dramsch JS, Lüthje M. Deep-learning seismic facies on state-of-the-art CNN architectures. In: SEG technical program expanded abstracts 2018. Houston: SEG Library; 2018.
- [19] Mosser L, Dubrulle O, Blunt MJ. Reconstruction of three-dimensional porous media using generative adversarial neural networks. *Phys Rev E* 2017;96(4):043309.
- [20] Laloy E, Héroult R, Jacques D, Linde N. Training-image based geostatistical inversion using a spatial generative adversarial neural network. *Water Resour Res* 2018;54(1):381–406.
- [21] Mousavi SM, Horton SP, Langston CA, Samei B. Seismic features and automatic discrimination of deep and shallow induced-microearthquakes using neural network and logistic regression. *Geophys J Int* 2016;207(1):29–46.
- [22] He M, Li Q, Li X. Injection-induced seismic risk management using machine learning methodology—a perspective study. *Front Earth Sci* 2020;8:227.
- [23] Mahani AB, Schultz R, Kao H, Walker D, Johnson J, Salas C. Fluid injection and seismic activity in the northern Montney play, British Columbia, Canada, with special reference to the 17 August 2015 M w 4.6 induced earthquake. *Bull Seismol Soc Am* 2017;107(2):542–52.
- [24] Ellsworth WL. Injection-induced earthquakes. *Science* 2013;341(6142):1225942.
- [25] Atkinson GM, Eaton DW, Ghofrani H, Walker D, Cheadle B, Schultz R, et al. Hydraulic fracturing and seismicity in the western Canada sedimentary basin. *Seismol Res Lett* 2016;87(3):631–47.
- [26] Gharbi RB, Elsharkawy AM. Universal neural network based model for estimating the PVT properties of crude oil systems. In: Proceedings of the SPE Asia Pacific Oil and Gas Conference and Exhibition; 1997 Apr 14–16; Kuala Lumpur, Malaysia. Richardson: OnePetro; 1997.
- [27] Osman EA, Abdel-Wahhab OA, Al-Marhoun MA. Prediction of oil PVT properties using neural networks. In: Proceedings of the SPE Middle East Oil Show; 2001 Mar 17–20; Manama, Bahrain. Richardson: OnePetro; 2001.
- [28] Wang K, Luo J, Wei Y, Wu K, Li J, Chen Z. Practical application of machine learning on fast phase equilibrium calculations in compositional reservoir simulations. *J Comput Phys* 2020;401:109013.
- [29] Helmy T, Fatai A. Hybrid computational intelligence models for porosity and permeability prediction of petroleum reservoirs. *Int J Comput Intell Appl* 2010;9(4):313–37.
- [30] Patel AK, Chatterjee S. Computer vision-based limestone rock-type classification using probabilistic neural network. *Geosci Front* 2016;7(1):53–60.
- [31] An P. The effect of random noise in lateral reservoir characterization using feed-forward neural networks. In: SEG technical program expanded abstracts 1994. Houston: SEG Library; 1994.
- [32] An P, Moon WM, Kalantzis F. Reservoir characterization using seismic waveform and feedforward neural networks. *Geophysics* 2001;66(5):1450–6.
- [33] Jamialahmadi M, Javadpour FG. Relationship of permeability, porosity and depth using an artificial neural network. *J Petrol Sci Eng* 2000;26(1–4):235–9.
- [34] Wang B, Sharma J, Chen J, Persaud P. Ensemble machine learning assisted reservoir characterization using field production data—an offshore field case study. *Energies* 2021;14(4):1052.
- [35] Liu X, Ge Q, Chen X, Li J, Chen Y. Extreme learning machine for multivariate reservoir characterization. *J Petrol Sci Eng* 2021;205:108869.
- [36] Chen Z. Reservoir simulation: mathematical techniques in oil recovery. In: CBMS-NSF regional conference series in applied mathematics. Philadelphia: Siam; 2007.

- [37] Chen Z, Huan G, Ma Y. Computational methods for multiphase flows in porous media. Philadelphia: Siam; 2006.
- [38] Amirian E, Chen Z. Cognitive data-driven proxy modeling for performance forecasting of waterflooding process. *Glob J Technol Optim* 2017;08(01):1–9.
- [39] Amirian E, Dejam M, Chen Z. Performance forecasting for polymer flooding in heavy oil reservoirs. *Fuel* 2018;216:83–100.
- [40] Huang Z, Chen Z. Comparison of different machine learning algorithms for predicting the SAGD production performance. *J Petrol Sci Eng* 2021;202:108559.
- [41] Dang C, Nghiem L, Fedutenko E, Gorucu SE, Yang C, Mirzabozorg A, et al. AI based mechanistic modeling and probabilistic forecasting of hybrid low salinity chemical flooding. *Fuel* 2020;261:116445.
- [42] Ng CSW, Ghahfarokhi AJ, Amar MN. Well production forecast in Volve field: application of rigorous machine learning techniques and metaheuristic algorithm. *J Petrol Sci Eng* 2022;208(Pt B):109468.
- [43] Hui G, Chen S, He Y, Wang H, Gu F. Machine learning-based production forecast for shale gas in unconventional reservoirs via integration of geological and operational factors. *J Nat Gas Sci Eng* 2021;94:104045.
- [44] Tadjer A, Hong A, Bratvold RB. Machine learning based decline curve analysis for short-term oil production forecast. *Energy Explor Exploit* 2021;39(5):1747–69.
- [45] Ahmad MA, Chen Z. Machine learning models to predict bottom hole pressure in multi-phase flow in vertical oil production wells. *Can J Chem Eng* 2019;97(11):2928–40.
- [46] Lee K, Lim J, Yoon D, Jung H. Prediction of shale-gas production at Duvernay Formation using deep-learning algorithm. *SPE J* 2019;24(06):2423–37.
- [47] Popa A, Connel S. Optimizing horizontal well placement through stratigraphic performance prediction using artificial intelligence. In: Proceedings of the SPE Annual Technical Conference and Exhibition; 2019 Sep 30–Oct 2; Calgary, AB, Canada. Richardson: OnePetro; 2019.
- [48] Mohaghegh SD. Mapping the natural fracture network in Utica shale using artificial intelligence (AI). In: Proceedings of the SPE/AAPG/SEG Unconventional Resources Technology Conference; 2017 Jul 24–26; Austin, TX, USA. Houston: SEG Library; 2017.
- [49] He Q, Zhong Z, Alabboodi M, Wang G. Artificial intelligence assisted hydraulic fracturing design in shale gas reservoir. In: Proceedings of the SPE Eastern Regional Meeting; 2019 Oct 15–17; Charleston, WV, USA. Richardson: OnePetro; 2019.
- [50] Sun Z, Wang L, Zhou J, Wang C. A new method for determining the hydraulic aperture of rough rock fractures using the support vector regression. *Eng Geol* 2020;271:105618.
- [51] Yang S, McBride P, Kherroubi J, He A, Le Nir I, Quesada D, et al. An efficient workflow for geological characterization in unconventional reservoirs from a new through-the-bit logging electrical micro-imaging tool. In: Proceedings of the 2018 AAPG International Conference and Exhibition; 2018 Nov 4–11; Cape Town, South Africa; 2018.
- [52] Wang LK, Sun AY. Well spacing optimization for Permian basin based on integrated hydraulic fracturing, reservoir simulation and machine learning study. In: Proceedings of the SPE/AAPG/SEG Unconventional Resources Technology Conference; 2020 Jul 20–22; online. Richardson: OnePetro; 2020.
- [53] Bangi MSF, Kwon JSI. Deep reinforcement learning control of hydraulic fracturing. *Comput Chem Eng* 2021;154:107489.
- [54] Duplyakov VM, Morozov AD, Popkov DO, Shel EV, Vainshtein AL, Burnaev EV, et al. Data-driven model for hydraulic fracturing design optimization. Part II: inverse problem. *J Petrol Sci Eng* 2021;208(Pt A):109303.
- [55] Chaikine IA, Gates ID. A machine learning model for predicting multi-stage horizontal well production. *J Petrol Sci Eng* 2021;198:108133.
- [56] Han D, Kwon S. Application of machine learning method of data-driven deep learning model to predict well production rate in the shale gas reservoirs. *Energies* 2021;14(12):3629.
- [57] Mohaghegh SD, inventor; Mohaghegh SD, assignee. System and method providing real-time assistance to drilling operation. United States patent US20150300151. 2015 Oct 22.
- [58] Unrau S, Torrione P, Hibbard M, Smith R, Olesen L, Watson J. Machine learning algorithms applied to detection of well control events. In: Proceedings of the SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition; 2017 Apr 24–27; Dammam, Saudi Arabia. Richardson: OnePetro; 2017.
- [59] Pollock J, Stoecker-Sylvia Z, Veedu V, Panchal N, Elshahawi H. Machine learning for improved directional drilling. In: Proceedings of the Offshore Technology Conference; 2018 Apr 30–May 3; Houston, TX, USA. Richardson: OnePetro; 2018.
- [60] Zhao J, Shen Y, Chen W, Zhang Z, Johnston S. Machine learning-based trigger detection of drilling events based on drilling data. In: Proceedings of the SPE Eastern Regional Meeting; 2017 Oct 4–6; Lexington, KY, USA. Richardson: OnePetro; 2017.
- [61] Hegde C, Gray KE. Use of machine learning and data analytics to increase drilling efficiency for nearby wells. *J Nat Gas Sci Eng* 2017;40:327–35.
- [62] Goebel T, Molina RV, Vilalta R, Gupta KD, inventors; Landmark Graphics Corp., assignee. Method and system for predicting a drill string stuck pipe event. United States patent US8752648. 2014 Sep 10.
- [63] Dursun S, Tuna T, Duman K, Kellogg RW, inventors; Landmark Graphics Corp., assignee. Real-time risk prediction during drilling operations. United States patent US15/024, 575. 2015 Apr 30.
- [64] Castifeira D, Toronyi R, Saleri N. Machine learning and natural language processing for automated analysis of drilling and completion data. In: Proceedings of the SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition; 2018 Apr 23–26; Dammam, Saudi Arabia. Richardson: OnePetro; 2018.
- [65] Bhattacharya S. A brief review of popular machine learning algorithms in geosciences. In: Briefs in Petroleum Geoscience & Engineering. Cham: Springer; 2021.
- [66] Abraham RJ, See LM, Solomatine DP. Practical hydroinformatics: computational intelligence and technological developments in water applications. New York: Springer Science & Business Media; 2008.
- [67] Li Z, Kovachki N, Azizzadenesheli K, Liu B, Bhattacharya K, Stuart A, Anandkumar A. Fourier neural operator for parametric partial differential equations. 2020. arXiv:2010.08895.
- [68] Noshi C. A brief survey of text mining applications for the oil and gas industry. International Petroleum Technology Conference; 2019 Mar 26–28; Beijing, China. Richardson: OnePetro; 2019.
- [69] Brestoff NE, inventor; Intraspexion LLC assignee. Using classified text and deep learning algorithms to identify risk and provide early warning. United States patent US9552548B1. 2017 Jan 24.
- [70] Lv G, Zheng C, Zhang L. Text information retrieval based on concept semantic similarity. In: Proceedings of the 2009 Fifth International Conference on Semantics; 2009 Oct 12–14; Zhuhai, China. Piscataway: IEEE Publisher; 2009.
- [71] Arumugam S, Rajan S, Gupta S. Augmented text mining for daily drilling reports using topic modeling and ontology. In: Proceedings of the SPE Western Regional Meeting; 2017 Apr 23–27; Bakersfield, CA, USA. Richardson: OnePetro; 2017.