International Journal of Business, Management and Visuals (IJBMV), ISSN: 3006-2705
Volume 5, Issue 2, July-December, 2022, Available online at: https://ijbmv.com

Machine learning in the petroleum and gas exploration phase current and future trends

Dipak Kumar Banerjee[1], Ashok Kumar[2], Kuldeep Sharma[3]
[1,3]Welspun Tubular Inc., Frazier Pike, Lr-72206 USA, Production Planning-NDT
[2]Welspun Corp Ltd., Anjar, Gujarat, India-Welding

ABSTRACT

We assert that artificial intelligence is poised to significantly impact the oil and gas industry, particularly in the upstream sector, which is the most expensive and high-risk area. Our analysis of AI applications reveals clear trends in the development of powerful tools that accelerate processes and effectively reduce risks. We emphasize the importance of various AI methods and the critical need for data availability. Additionally, we address the non-technical barriers hindering AI adoption, including data-related challenges and workforce limitations. Furthermore, we outline three definitive future scenarios for the transformative role of AI in the oil and gas sector over the next 5, 10, and 20 years. Data generation in the oil and gas industry is constant and substantial, making effective recording and utilization essential. Decision-making driven by predictive and inferential data analytics empowers organizations to make swift and accurate choices. Despite existing challenges, the adoption of data analytics is on the rise and is transforming the sector. Recent advancements in Artificial Intelligence (AI) and Machine Learning (ML) are revolutionizing the way complex problems are addressed and production is optimized. By leveraging both historical and real-time data from oil and gas wells, companies can significantly enhance their operational efficiency. The industry has embraced various analytical modeling techniques that facilitate decisive, data-driven strategies.

This paper decisively reviews the recent breakthroughs in AI and ML applications for data exploitation, spanning from crude oil exploration to product distribution. It also outlines the promising future of these technologies within the industry. The findings provide a clear framework for selecting the most effective technologies to manage and harness the vast data generated by the oil and gas sector.

Keywords: Artificial Intelligence, Machine learning, Upstream, Oil and gas industry, Petroleum exploration.

INTRODUCTION

The oil and gas sector encompasses a wide range of critical activities, including the exploration of oil fields, reservoir engineering, drilling, and production engineering. These industries fundamentally depend on oil and gas as primary sources for the production of essential chemicals, such as pharmaceuticals, solvents, fertilizers, pesticides, and plastics (Anderson, 2017). As the cost of fossil fuels continues to rise, companies in this sector must innovate and optimize their operations to enhance efficiency and fully leverage their existing capabilities. The reality is that many oil fields are now mature, often producing more water than oil due to issues like water influx nearshore, channeling, coning, or water breakthrough. This situation makes it economically unfeasible to extract petroleum from these formations. Moreover, the volatile nature of oil prices has led most oil and gas firms to prioritize immediate financial stability over investing in expensive engineering solutions and equipment. To effectively tackle these challenges, the adoption of Inflow Control Devices (ICD) or Inflow Control Valves (ICV), along with advanced downhole sensor systems, presents a powerful strategy for enhancing efficiency and productivity. These technologies will maximize the extraction of resources, addressing key operational demands head-on. Achieving effective control in major oilfields requires decisive and timely decision-making amid ongoing challenges. The concept of a "Smart Oilfield" is not merely a trend; it represents a necessary evolution in the industry. It entails developing a robust technological infrastructure that includes digitizing instrumentation systems and fostering a collaborative network to optimize production processes. The impact of digital technology on business and society is undeniable. Digital transformation is rightfully labelled the "fourth industrial revolution," driven by the integration of cutting-edge technologies that blur the boundaries between the physical, digital, and biological worlds, including artificial intelligence, robotics, and autonomous vehicles.

Artificial Intelligence (AI) technologies are making waves due to their rapid response capabilities and exceptional ability to generalize. Machine learning, in particular, offers substantial potential for transforming traditional reservoir engineering methodologies across a multitude of complex challenges Research in this field has established the effectiveness of sophisticated machine learning algorithms, including Fuzzy Logic (FL), Artificial Neural Networks (ANN), Support Vector Machines (SVM), Response Surface Models (RSM), and various classification and regression tools . The majority of these algorithms operate within the realm of supervised learning. In traditional reservoir engineering applications, evolutionary optimization techniques, such as Genetic Algorithms (GA) and Particle Swarm Optimization (PSO), are indispensable. To solve inverse problems decisively, ongoing research must prioritize the development of analytical workflows that seamlessly integrate forward-looking AI models with reverse-looking models. A notable example is the work of, who successfully developed AI-assisted workflows on a unified platform, utilizing forward-looking Gaussian proxy designs, Bayesian optimization, and high-fidelity numerical models. This approach effectively addressed specific challenges in a coal seam degasification program based on historical data. Bayesian optimization is particularly valuable as it can identify multiple solutions for reservoir characteristics that align with available field information .[1][2]Digitalization, digital twins (DT), Industry 4.0, and the industrial Internet of Things (IIoT) are revolutionizing the oil and gas (O&G) sector by effectively addressing the risks associated with exploration, drilling, and distribution. Faced with stringent regulatory pressures, a looming skills gap due to retiring employees, and prolonged low oil prices, O&G companies must innovate decisively to enhance productivity, reduce health, safety, and environmental (HSE) risks, and ensure regulatory compliance. The rapid advancement of technology is reshaping business models and creating significant new revenue-generating opportunities. Innovations in cloud computing, artificial intelligence, the Internet of Things (IoT), and blockchain technology are enabling robust cyber-physical integration. This integration allows companies to collect and analyze data effectively to inform decisions and optimize operations through the implementation of the DT concept. The Gartner Group has rightfully identified DT as a leading strategic technology trend. This concept has already been successfully applied across various sectors, including manufacturing, healthcare, and aviation. As the O&G industry starts to embrace digital technologies, it is essential that they move beyond a bottom-up approach. The current fragmented implementation limits the recognition of the full benefits of digitalization and DT.A comprehensive understanding of DT technology and the current state of research, as well as the challenges ahead, is essential for its successful deployment in O&G. Key application areas identified through a thorough literature review include asset integrity monitoring and project planning, while significant challenges such as cyber security and the lack of standardization must be addressed head-on. Research indicates that the United States is at the forefront of DT-related publications in the O&G sector, with a notable surge in publications since 2017. Despite this progress, many studies still focus on theoretical frameworks rather than practical applications, which highlights that DT implementation is still in its nascent stages.

This article is structured to provide valuable insights: Section II outlines the literature survey methodology, Sections III and IV review DT and its research status, Section V examines the opportunities and challenges, and the conclusion synthesizes the critical findings. The path forward is clear; the O&G industry must fully leverage DT technology to thrive in the modern landscape.[2][3]Data is the cornerstone of successful machine learning pipelines, and the oil and gas sector has unmatched access to vast amounts of it. Subsea sensors capture data in real-time, often every few seconds, while seismic surveys can generate over 6 terabytes of data daily. With significant on-premise and cloud computing capabilities, companies can conduct large-scale experiments and tap into expert knowledge to enhance algorithm training and improve prediction accuracy. However, the industry's risk-averse culture stifles technology adoption, and ineffective innovation management practices persist. Many companies lack a framework to manage machine learning and analytics initiatives effectively. While software companies thrive on agile methodologies, the oil and gas sector remains mired in slower, hierarchical processes that hinder innovation. To accelerate advancements, the industry must leverage hardware improvements, including Graphics Processing Units (GPUs) and Field-Programmable Gate Arrays (FPGAs), which can greatly enhance training speeds. Additionally, transfer learning—transferring knowledge from one task to optimize another—should be embraced to overcome the scarcity of well-annotated datasets. Finally, adopting Continuous Integration/Continuous Deployment (CI/CD) practices in machine learning is critical. A robust CI/CD system should ensure reliable, reproducible pipelines with essential features for monitoring and version control, addressing the challenges posed by concept drift. The industry must act decisively to modernize its approach and fully harness the power of data.[3]

MATERIAL & METHODS

To effectively prevent damage during drilling operations for new offshore oil and gas wells, a semi-submersible drilling unit must maintain its position directly above the wellhead. This requirement is critical, particularly in shallow waters, where even minor shifts can significantly alter the riser angles (the pipe connecting the platform to the subsea drilling system). Exceeding the physical inclination limits can lead to serious damage to key components, including the wellhead, Blowout Preventer (BOP), and Lower Marine Riser Package (LMRP).The platform's positioning is autonomously and rigorously maintained using a robust Dynamic Positioning (DP) system, which controls a network of thrusters without the need for a mooring system. This DP system receives precise input from a position reference system that incorporates Differential Global Positioning System (DGPS) and Hydroacoustic Position Reference (HPR), along with environmental sensors, a gyrocompass, radar, and an inclinometer. A Dynamic Positioning Operator (DPO) stationed in the Marine Control Room (MCR) vigilantly monitors the DP panels and screens, standing ready to execute emergency procedures as necessary. Loss of platform position can occur for various reasons. In this case study, we assert that the platform's thrusters may inadvertently propel the unit in the wrong direction, resulting in a “drive-off” scenario. If the rig shifts to an offset position, designated alarms will activate, requiring the DPO to immediately halt the drive-off by deactivating the thrusters and initiating the manual Emergency Disconnect Sequence (EDS) to disconnect the riser from the BOP. Should the manual EDS fail, the automatic EDS will engage at the ultimate position limit to ensure a safe disconnection They conducted an in-depth analysis of drive-off scenarios, modeling each safety barrier outlined in the event tree of with comprehensive hierarchical structures that encompass technical, operational, and organizational systems. She established relevant indicators to rigorously assess the performance of these systems and their corresponding barriers. A total of 50 indicator categories were defined, with values meticulously collected and translated onto a criticality scale ranging from 1 to 6. These indicator trends were simulated over a 30-year period, inspired by the typical bathtub curve for technical elements and informed expert judgment for the remaining elements. As demonstrated these indicator values can and should be aggregated based on their relative weights and hierarchical barrier models, allowing for dynamic updates of barrier failure data.[4][8]The utilization of statistical models for Rate of Penetration (ROP) prediction presents both similarities and distinct advantages over traditional approaches. While both require pre-selection of models based on drilling variables, statistical models do not attempt to capture the intricacies of drill bit operation and the interactions between rock formations and the bit, as conventional models do. The 53 studies analyzed demonstrate the application of machine learning (ML) techniques in ROP prediction, which can be confidently categorized into five primary methods: artificial neural networks (ANN), support vector machines (SVM), fuzzy logic systems, neuro-fuzzy systems, and ensemble models.

Artificial neural networks (ANN) vary in their structures, activation functions, and training methodologies. A prevalent structure is the feedforward neural network with a multilayer perceptron (MLP), which typically consists of one or two layers. The first layer effectively connects inputs to hidden units, while the second layer links hidden units directly to outputs, utilizing adaptive weights that include necessary bias terms.

Ensemble models excel by combining predictions from multiple "weak learners." These can be categorized as homogeneous, consisting of the same type of learners, or heterogeneous, incorporating various techniques. Notably, some experts assert that traditional ROP models should be employed alongside ML models, particularly when datasets are limited. Traditional models require less data yet still produce reliable predictions. Integrating these approaches into a unified model is entirely feasible through aggregation functions, which are fundamental to ensemble methods. Importantly, this synergistic combination warrants further exploration, as it has not yet been adequately addressed in existing literature.[8]

RESULTS

The primary objective of machine learning is to develop a robust model that excels across various dimensions. However, real-world conditions often present challenges that deviate from this ideal. Ensemble learning effectively addresses this issue by integrating multiple models to create a powerful and comprehensive supervised model. The principle behind ensemble learning is clear: when one weak classifier errs, others can rectify that mistake. We present a groundbreaking hybrid approach that combines Multilayer Perceptron (MLP), Support Vector Regression (SVR), and CatBoost to accurately predict energy consumption. The innovative structure of our proposed hybrid model is depicted in Figure 3. This model not only trains the CatBoost, MLP, and SVR algorithms but also effectively synthesizes their forecasting results to ensure superior accuracy.[5][6]The successful implementation of data-driven methods in the petroleum industry requires a strong understanding of both petroleum engineering processes and the physics-based conventional techniques, along with proficiency in traditional statistics, data mining, artificial intelligence, and machine learning. These methods begin with a data-centric approach to identify issues and develop solutions. While data-driven methods can provide effective solutions for challenging and complex processes that are difficult to define using existing conventional methods, there remains skepticism within the industry regarding their use. This skepticism is closely linked to the delicate and sensitive nature of the processes involved and the handling of data. Proper organization and refinement of data are crucial components of an efficient data-driven process.

Data-driven methods offer significant advantages over conventional methods under certain conditions. However, many industry professionals still have a vague understanding of these methods.[6][7]Oil spills in seas and oceans are a critical source of maritime pollution, driven primarily by human activity and rising oil demand. Their impacts on aquatic ecosystems, wildlife, tourism, aquaculture, and coastal economies are severe and undeniable. Continuous monitoring and prompt intervention are not just important—they are essential for mitigating these environmental crises. Remote monitoring capabilities are indispensable for the protection of marine biodiversity and habitats. The last decade has seen significant advancements in oil spill detection, fueled by increased access to remotely sensed data, enhanced computational power, cloud computing, and cutting-edge machine learning algorithms. Satellite and airborne remote sensing techniques, utilizing a range of sensors such as multispectral, hyperspectral, thermal, and microwave, have proven effective in detecting and estimating the thickness of oil spills. Notably, microwave satellite-based synthetic aperture radar (SAR) stands out for its performance under various weather conditions. The growing reliance on satellite-based multispectral data empowers us to differentiate oil spills from lookalikes, yet the use of ultraviolet and laser fluorosensors remains underutilized. The accuracy of oil spill detection is critically impacted by the similarities between oil spills and other natural or manmade features. Integrating various feature categories—such as statistical, geometric, and texture features—is vital to enhancing classification accuracy. Despite this, many studies still depend on manual feature extraction based on analysts' experience, while only a limited number leverage advanced feature selection techniques, which are essential for improving classification reliability. It is imperative to evaluate the efficiency of features extracted from remote sensing images.

The acquisition of high-quality training samples is a fundamental requirement for effective machine learning classification. The question of how many samples are necessary for dependable results remains unresolved. Collecting accurately labeled samples is a significant challenge due to similarities with lookalikes, which can easily mislead even the most experienced analysts, resulting in unacceptable false positives and negatives in detection efforts. Immediate action is needed to enhance the efficacy of oil spill detection systems for the protection of our marine environments.[7]

CONCLUSION

This paper has explored various approaches to combining data and utilizing machine learning for making predictions in the context of Industry 4.0. It has examined three primary types of predictive models: descriptive, predictive, and prescriptive, each designed with a specific purpose in mind. To enhance understanding of the current state of research in this area, a comparison of these models has been presented. Based on the literature review in earlier sections, it is evident that there is substantial interest in employing data fusion and analysis for industrial predictions. In fact, nearly every new design for Industry 4.0 systems highlights the significance of predictive models for the efficient operation of the targeted industrial assets. By applying data fusion methods and machine learning algorithms, it is possible to leverage all available data effectively.

REFERENCES

[1]. Anirbid Sircar, Kriti Yadav, Kamakshi Rayavarapu, Namrata Bist, Hemangi Oza-Petroleum Research, Volume 6, Issue 4, December 2021, Pages 379-391-Application of machine learning and artificial intelligence in oil and gas industry.

[2]. THUMEERA R. WANASINGHE1,(Member, IEEE), Leah Wroblewski1,3, Bui K. Petersen1,Raymond G.Gosine1,2,Lesley Anne James1,Oscar De Silva1, (Member, IEEE), George K. I. Mann 1, (Member, IEEE), And Peter J. Warrian, IEEE Access-Received April 29, 2020, accepted May 14, 2020, date of publication June 1, 2020, date of current version June 15, 2020.-Digital Twin for the Oil and Gas Industry: Overview, Research Trends, Opportunities, and Challenges.

[3]. Yasin Hajizadeh, Journal of Petroleum Science and Engineering 176 (2019) 661–663-Machine learning in oil and gas; a SWOT analysis approach

[4]. Nicola Paltrinieri, Louise Comfort, Genserik Reniers-Safety Science, Volume 118, October 2019, Pages 475-486,Learning about risk: Machine learning for risk assessment

[5]. Prince Waqas Khan, Yung-Cheol Byun, Sang-Joon Lee, Dong-Ho Kang, Jin-Young Kang, Hae-Su Park-Energies 2020, 13(18), 4870; https://doi.org/10.3390/en13184870-Machine Learning-Based Approach to Predict Energy Consumption of Renewable and Nonrenewable Power Sources.

[6]. Karthik Balaji; Minou Rabiei; Vural Suicmez; Celal Hakan Canbaz; Zinyat Agharzeyva; Suleyman Tek; Ummugul Bulut; Cenk Temizel- SPE-190812-MS

[7]. https://doi.org/10.2118/190812-MS Published: June 11 2018-Status of Data-Driven Methods and their Applications in Oil and Gas Industry

[8]. Rami Al-Ruzouq, Mohamed Barakat A.Gibril, Abdallah Shanableh, Abubakir Kais, Osman Hamed, Saeed Al-Mansoori, Mohamad Ali Khalil-Remote Sens. 2020, 12(20), 3338; https://doi.org/10.3390/rs12203338-13 October 2020-Sensors, Features, and Machine Learning for Oil Spill Detection and Monitoring: A Review

[9]. Luís Felipe F.M. Barbosaa, Andreas Nascimentob, Mauro Hugo Mathiase, João Andrade de Carvalho Jr.-Journal of Petroleum Science and Engineering 183 (2019) 106332-Machine learning in the petroleum and gas exploration phase current and future trends