ISSN 2616-5767 Vol.8, Issue 1: 24-31, DOI: 10.25236/AJETS.2025.080104

Applications of Big Data in Oil and Gas Industry of Asia - Taking China, India and Pakistan as Examples

Jianyu Miao[1,a,*], Rimsha Rahat[1,b], Khubaib Ubaid Tambe[1,c]

[1]Academy of Engineering, Peoples' Friendship University of Russia, Moscow, Russia
[a]jianyumiaocn@gmail.com, [b]rimsha.rahat125@gmail.com, [c]khubaibtambe564@gmail.com
[*]Corresponding author

Abstract: *The integration of big data analytics technologies in the oil and gas industries of China, Pakistan, and India is revolutionizing the industry, improving efficiency, enhancing safety, and supporting sustainable development in these resource-dependent economies. This paper demonstrates the unique applications of big data in each country's exploration, production, and distribution phases, highlighting how predictive analytics, real-time monitoring, and advanced simulations can optimize resource management. In China, extensive big data infrastructure supports smart oilfield development and automated production management, while Pakistan uses data analytics to achieve exploration precision in challenging terrain. As energy demand continues to grow, India uses big data to maximize extraction efficiency and streamline regulatory compliance. By analyzing data extracted from studies in these countries, this paper explores the challenges faced, such as data infrastructure gaps and skills shortages, and the strategic approaches each country has taken to address these challenges. Ultimately, big data applications in these countries demonstrate their critical role in promoting energy exploration, production efficiency, and environmental responsibility, positioning big data as a key tool for the future of the oil and gas industry in South and East Asia.*

Keywords: *Oil Exploration, Big Data, Intelligent Oilfield development, Enhanced Oil Recovery, Environmental Protection*

1. Introduction

The application of big data technologies in the oil and gas industries of China, Pakistan, and India is revolutionizing operational efficiency, safety, and environmental sustainability. In China, advances in automation and data management enable sophisticated platforms such as PetroChina's DreamCloud to support real-time monitoring and data lake systems that drive insights during the exploration and production phases. Pakistan's oil and gas industry leverages large amounts of seismic data and advanced analytics to optimize asset performance and address environmental issues. Big data enables better decision-making through sensor-based monitoring and predictive maintenance, delivering environmental and economic benefits. In India, big data combined with the Internet of Things, machine learning, and edge computing facilitates real-time analytics and improves safety measures. Technologies such as digital twins enable predictive maintenance, minimize downtime, and enhance resource management. Despite the high costs and complex infrastructure required, these countries are investing in digital transformation, demonstrating that big data plays a vital role in the future sustainability and growth of their oil industries. This interconnected approach promises to create an optimized, efficient, and environmentally responsible energy industry. The adoption of big data in these countries has not only optimized oil and gas operations, but has also fostered collaborative innovation with technology companies such as Huawei and PwC. This collaboration accelerates the transition to the "smart oil field," where artificial intelligence, cloud computing and edge technologies streamline processes and enhance resiliency. As industries overcome challenges related to data security and infrastructure, big data will redefine energy production standards, enabling countries to achieve sustainable growth, cost-effectiveness and security. These efforts underscore a future where data-driven insights enable proactive approaches to managing resources and addressing environmental impacts in a changing global energy landscape.

2. The development of big data in Pakistan

The development of big data in Pakistan has grown significantly over the past decade, driven by

ISSN 2616-5767 Vol.8, Issue 1: 24-31, DOI: 10.25236/AJETS.2025.080104

technological advancements, increased digital connectivity, and a booming tech industry. Initially, big data adoption was limited, mostly limited to the telecommunications and banking industries. However, as internet penetration and mobile phone usage rise, more and more industries, including healthcare, e-commerce, and agriculture, are beginning to explore data-driven solutions.

The "Digital Pakistan" policy is an important means for the Pakistani government to develop big data technology. It aims to further accelerate the use of big data technology in decision-making assistance in various industries in Pakistan and encourage data-driven governance and smart city projects. In addition, cooperation between Pakistani technology companies and international organizations has introduced advanced analytics and artificial intelligence (AI) into the Pakistani market and cultivated a data science ecosystem. Although the government continues to improve its own technological tone, challenges remain, especially in data privacy and infrastructure, so the Pakistani government needs to steadily develop its big data capabilities to promote economic growth and innovation in various industries. In addition, how to integrate the ANP model into the oil and gas industry to improve production efficiency and take advantage of enhanced technology. Pakistan has also launched a fruitful and ambitious project with the People's Republic of China, the China-Pakistan Economic Corridor (CPEC), which has helped the Pakistani government develop its technology and economy.

2.1. ANP Modeling Technique.

Currently, many research projects use both AHP (Analytical Hierarchy Process) and ANP (Analytical Network Process) to prioritize ICT policies to maximize the efficiency of the industry. Many countries have used the ANP model in their different case studies, but more research is needed if this model is to be applied in the case study of Pakistan.

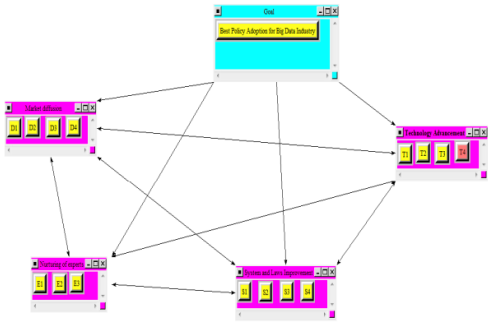

Figure 1: ANP general correlation for the study.

Figure 1 shows the general relevance of the ANP technique that has been used in the study of big data technology in Pakistan [1] . The ANP model is widely used for its multi-criteria decision-making solution regarding the number of judgment elicitations and rank reversal advantages. The technique has proven to be valuable to multiple parties simultaneously as it brings about a solution that everyone agrees on [2] . from the article that the ANP model can significantly enhance the decision-making capabilities of Pakistan's oil and gas industry by providing a structured, data-driven approach to address complex, interdependent challenges. The ANP model is an extension of the Analytical Hierarchy Process (AHP) that incorporates feedback and interdependence into decision-making criteria, making it particularly suitable for the oil and gas industry with complex supply chains, regulatory factors and environmental considerations. At the same time, ANP can provide assistance in areas such as project selection, resource allocation, risk assessment and supplier evaluation. For example, it can help prioritize exploration sites by assessing factors such as geological potential, environmental impact and political risk, considering how each factor affects the others. ANP can also support supply chain decisions, weighing cost, reliability and geopolitical stability among supplier and logistics options.

The integration of ANP will be particularly valuable in balancing Pakistan's energy needs with sustainable development goals and regulatory compliance, ultimately supporting smarter, more transparent and agile decision-making processes across the industry.

ISSN 2616-5767 Vol.8, Issue 1: 24-31, DOI: 10.25236/AJETS.2025.080104

3. Big Data Applications in India's Oil and Gas Industry

In India's oil and gas industry, big data analytics plays a vital role in improving operational efficiency, safety, and environmental sustainability. Applications range from optimizing production processes to real-time monitoring of equipment and reducing maintenance costs. Technologies such as machine learning, IoT, and digital twins enable predictive maintenance, allowing companies to predict and resolve equipment failures, minimize downtime, and increase productivity. In addition, big data assists asset management and operational decisions by using sensor data on drilling rigs and pipelines, ensuring safety and environmental compliance by monitoring emissions and hazardous conditions.

However, implementing big data analytics faces many challenges, such as high costs, data quality issues, and complex data transfer from remote sites. These difficulties, coupled with the need for advanced IT infrastructure and network security, make the full adoption of big data analytics a gradual process. However, as more companies invest in these technologies, the role of big data in the Indian oil and gas industry is expected to grow significantly.

3.1. Indian Oil and Gas Industry Leverages Edge Computing for Digital Transformation

Edge computing [3] offers transformative benefits by transforming raw data into actionable insights in the field, helping to optimize systems, enhance workflows, and improve operational efficiency. By collecting and analyzing data in real time from a variety of sensors, it enables on-site diagnostics and predictive assessments, allowing for immediate responses to any anomalies detected in the workflow. Additionally, edge computing helps improve workplace safety by collecting health data from medical sensors or wearable devices used by workers in the field. This enables AI/ML models to monitor health conditions and immediately signal when medical intervention is needed. Near-edge devices are useful for asset tracking, environmental monitoring, predicting equipment failures, and effective asset coordination in oil exploration and development.

Supervisory Control and Data Acquisition (SCADA) systems collect and analyze data in real-time from large amounts of complex machinery and equipment located in remote and rugged terrain in the oil and gas industry. However, SCADA systems rely heavily on the cloud and local DCs when processing and storing this data. In this context, edge computing can play a huge role as it leverages the distribution of edge devices at or near the data source, bringing the utility of cloud computing close to the data generation source. Edge computing can be used to monitor, store, analyze, and run AI and ML models and other critical applications by the SCADA system in the cloud using local edge devices.

3.2. How PwC's technology consulting supports India's oil and gas industry

Globally, capital-intensive industries such as oil and gas are increasingly adopting digital technologies such as edge computing. Large oil companies use edge computing to monitor equipment, analyze data to reduce costs, and improve productivity by integrating these technologies with existing legacy systems. This integration allows real-time data processing and advanced analytics, providing early predictions to support faster and more informed decision making.

A PwC study [4] shows that leaders in the oil and gas industry are focusing on data-driven technologies, including manufacturing execution systems (MES), cloud computing, and Internet of Things (IoT) solutions. These technologies enhance data management, enable remote monitoring, and leverage machine learning for predictive maintenance, making operations more efficient and adaptable to future needs.

PwC's Emerging Technologies Advisory team helps organizations address Industry 4.0 by performing readiness analyses, assessing technology maturity, aligning technology with business strategy, and managing risk. For clients adopting edge computing, PwC creates client value by developing a roadmap for specific use cases based on the client's unique infrastructure needs.

Many oil and gas companies face challenges in effectively leveraging data from traditional IT systems, especially in harsh, operational technology (OT)-intensive environments. Edge computing facilitates IT-OT convergence, combining these systems to overcome issues of resiliency, visibility, connectivity, and security in real time.

PwC's Technology Advisory Services also provides a structured readiness assessment that evaluates capabilities across multiple dimensions to provide a clear understanding of where improvements are needed. With a strategy-to-execution approach, PwC helps clients design, select and implement edge

ISSN 2616-5767 Vol.8, Issue 1: 24-31, DOI: 10.25236/AJETS.2025.080104

computing solutions to transform operating models and drive internal and external business growth.

4. The development and description of China's big data technology

The exploration and development technology of the oilfield industry is changing with each passing day, and the degree of automation and informatization is constantly improving. China's oil industry has entered the era of digitalization and intelligence. Along with this, its data volume has also rapidly broken through from MB level to TB level or even PB level, showing an exponential growth [5] . For upstream exploration and development, the intelligent platform for big data of oil and gas resource development can realize the rapid elastic scaling of massive oilfield data. Based on the distributed computing engine Spark, the π stream data processing system can be developed, which can use different deep learning algorithms to perform pipeline processing and analysis on oil and gas data, and complete the real-time monitoring of oil and gas field development through the oil and gas integrated control large screen. Through the oil and gas integrated control large screen, the real-time monitoring, early warning and display of oil and gas field development can be completed [6] , which can effectively evaluate and guide oilfield development plans, as shown in Figure 2,3 .

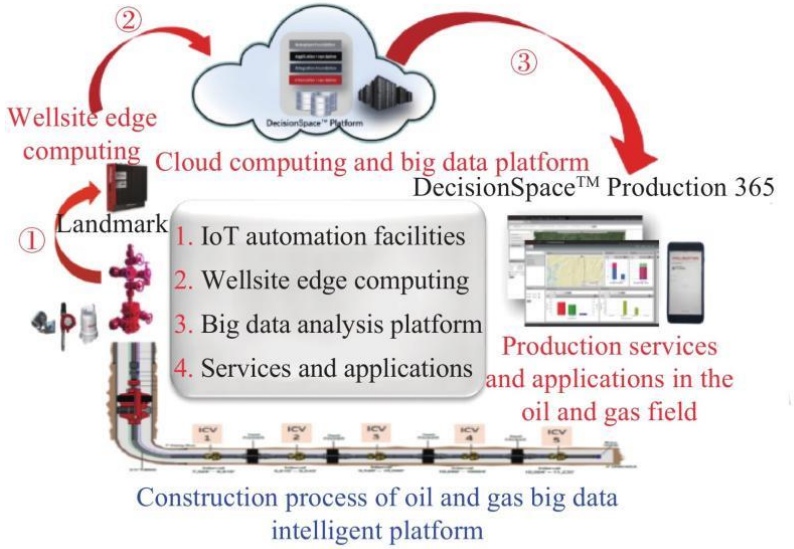

Figure 2: Construction of the intelligent platform for domestic oil and gas big data.

Figure 3: Example of the intelligent platform for domestic oil and gas big data.

The construction of the reservoir big data platform enables the oil industry to break the problem of "data islands" and extract and integrate data from various links such as exploration, drilling, completion,

ISSN 2616-5767 Vol.8, Issue 1: 24-31, DOI: 10.25236/AJETS.2025.080104

and production to create a reservoir data resource pool with good data quality, strong correlation, and high credibility. Of course, data resources are the cornerstone of the oil industry. If we compare it to a "weapon" in our hands, then artificial intelligence is the key to determining how powerful this "weapon" can be. As we all know, machine learning and deep learning have developed rapidly since the 21st century, and have gradually been combined with the oil industry to solve many complex oil engineering problems. It has become a hot topic for various scholars at home and abroad, such as the "Dream Cloud" platform released by China National Petroleum Corporation in November 2018, which is the first full platform in China's oil and gas industry. This is the first intelligent cloud platform in the domestic oil and gas industry to build a unified exploration and development data lake and a unified cloud platform for the entire platform. Relying on data lake and PaaS cloud platform technology, it has unified the data lake and unified management of 530,000 wells, 600 oil and gas reservoirs, 7,000 seismic zones, and 40,000 station libraries spanning more than 60 years, totaling 1.7PB, laying a solid data foundation for intelligent innovation of upstream business and forming the largest exploration and development data lake in China. The application of the collaborative research environment has achieved the expected results and has been applied in 156 exploration and research projects. The data preparation time has been shortened from 5 hours to less than 1 minute. Through "one-click" drawing, the drawings are automatically generated within a few seconds, realizing the transformation of exploration business research work from offline to online, from single person to collaborative, and from manual to automatic, and effectively optimizing the workflow. It has realized the transformation of exploration business research work from offline to online, from manual to collaborative, and from manual to automated, effectively optimizing the workflow, realizing the transformation of exploration business research results, and greatly improving work efficiency and decision-making level.

In addition, the Huawei Global Oil and Gas Summit held on July 15, 2020 showed that Huawei is committed to becoming a loyal partner in the digital transformation of oil and gas. It has used technologies such as 5G networks, big data, and cloud computing to solve the computing performance and data storage problems of PetroChina Daqing Oilfield, and has implemented the application of AI analysis in five major scenarios, including well logging oil and gas reservoir identification, seismic wave identification, and troubleshooting. Overall, although my country's oil and gas industry is still in the early stages of digital transformation, major oil and gas companies are actively building "smart oil fields" and working with data technology companies such as Huawei to enable the rapid development of digitalization and intelligence in the upstream and downstream of the oil and gas industry. I believe that this will be realized in the near future. I believe that in the near future, a brand new oil and gas industry ecosystem will be born.

4.1. China's big data applied to oil and gas resource development

As one of the typical representatives of old onshore oil fields in eastern China, Shengli Oilfield has accumulated a large amount of data resources from multiple sources and scales after 60 years of exploration and development. It has realized the real-time collection and storage of source data for earthquakes, geology, development dynamics, experimental analysis, oil production engineering, and the improvement of information construction level, and provided a strong software and hardware equipment foundation support for efficient operation and processing [7] . Shengli Oilfield Data Center currently stores 86 oil and gas fields, 4152 blocks/units, 8965 exploration wells, and 67241 development wells, totaling 760 million data and 5TB, with an average of more than 300,000 new data per day, providing data services for 151 applications and providing favorable conditions for the exploration and research of big data and artificial intelligence technologies. Next, the author will elaborate on the actual application of big data technology in the exploration and development stage, explore how big data technology can integrate the exploration and development business of old oil wells, and help oil fields to explore efficiently and develop efficiently. At the same time, the current adjustment situation will be analyzed and suggestions for the next development direction will be put forward.

4.2. Application of big data in oil and gas exploration

In view of the complex fracture systems, reservoir types and other geological characteristics of the exploration area, scientific researchers working in Shengli Oilfield actively explored the application research of big data and artificial intelligence technologies in the field of oil and gas exploration, improved and innovated the technologies required for several key work nodes of oil and gas intelligent modeling, such as layer extraction, remaining oil analysis and saturation prediction, and achieved a number of phased application results and experience.

ISSN 2616-5767 Vol.8, Issue 1: 24-31, DOI: 10.25236/AJETS.2025.080104

4.2.1. Automatic extraction of geological layers

For geological modeling, stratigraphic interpretation is a crucial step. Conventional stratigraphic extraction techniques rely on manual intervention, which results in heavy workload and insufficient interpretation results. Current automated or semi-automated structural extraction methods are usually based on phase tracking [8] , dip guidance [9], and waveform classification [10] to extract stratigraphic information from seismic data one by one. Most of these methods are based on local data and are difficult to extract data from noisy or complex geological structures. To this end, STARK.T introduces relative geological age bodies6 for stratigraphic extraction [11] . Through the processing of big data technology, the recognition target is changed from a single stratigraphic layer to all stratigraphic information, thereby obtaining useful stratigraphic information. In order to enhance the generalization ability of the network, the stratigraphic layers with known structural constraints are integrated to automatically extract a multi-task network model (see Fig. 4). First, a relative geological age body sample library with rich geological model orthogonal and geophysical orthogonal generation modes is established in combination with the structural characteristics of the exploration area; second, a structural automatic interpretation network model that can simultaneously output faults and relative geological age bodies is established; third, the loss function is improved in the matching of known stratigraphic constraint equations to enhance the generalization ability of the network model in deep learning. After matching, through continuous testing and optimization, the final horizon of the automatically extracted network model is obtained. The network model is used to extract high-precision horizon automatic interpretation results, and well-seismic calibration is used to quickly predict the seismic data corresponding to the relative geological age body.

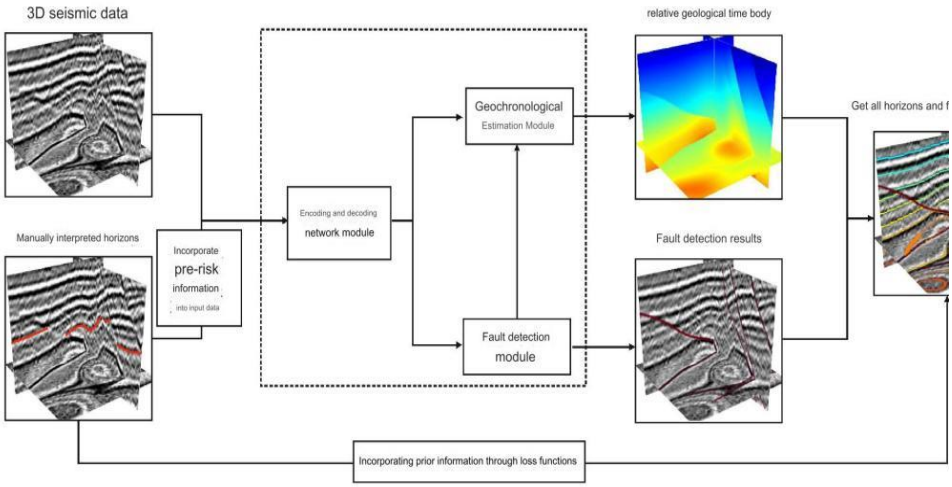

Figure 4: Multi-task network model for automatic horizon extraction integrating known structural restraints

4.2.2. Intelligent prediction of remaining oil saturation

When convolutional neural networks and other deep learning networks make predictions, the input and output parameters are independent of each other, that is, their output results only consider the current input, but not the input at other times, and there is no correlation in order. Therefore, recurrent neural networks came into being. They have a cyclic network structure inside, allowing information to be gradually transmitted outward. In addition to being related to the input of the current step, their output results also depend on the output of the previous step in the calculation of the current step, as if they have the ability to remember past data [12] . There are many time series-related parameters in the oil reservoir field, such as production, pressure, saturation, etc., which change over time. As shown in Figures 5 and 6 , Zhang et al. [13] used the LSTM deep learning model, a variant network of RNN, to learn the reservoir pressure field, water saturation field and historical production change data, predict the future residual oil saturation distribution, and compare it with the numerical simulation prediction results, which can effectively guide the secondary development of oil fields. Figure 6 shows the production and residual oil distribution prediction process based on LSTM neural network, which contains the LSTM neural network unit structure composed of forget gate, input gate and output gate, where tanh and σ are the corresponding activation functions in the network. Figure 5 shows the comparison of water saturation and pressure field prediction results. The black dots in the figure are well locations and corresponding well numbers.

ISSN 2616-5767 Vol.8, Issue 1: 24-31, DOI: 10.25236/AJETS.2025.080104

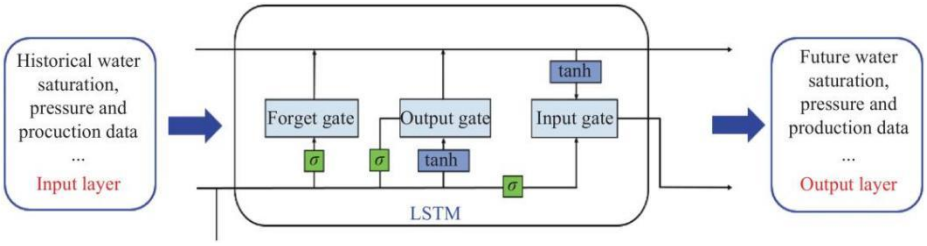

Figure 5: Prediction process of production data and remaining oil distribution based on LSTM neural network

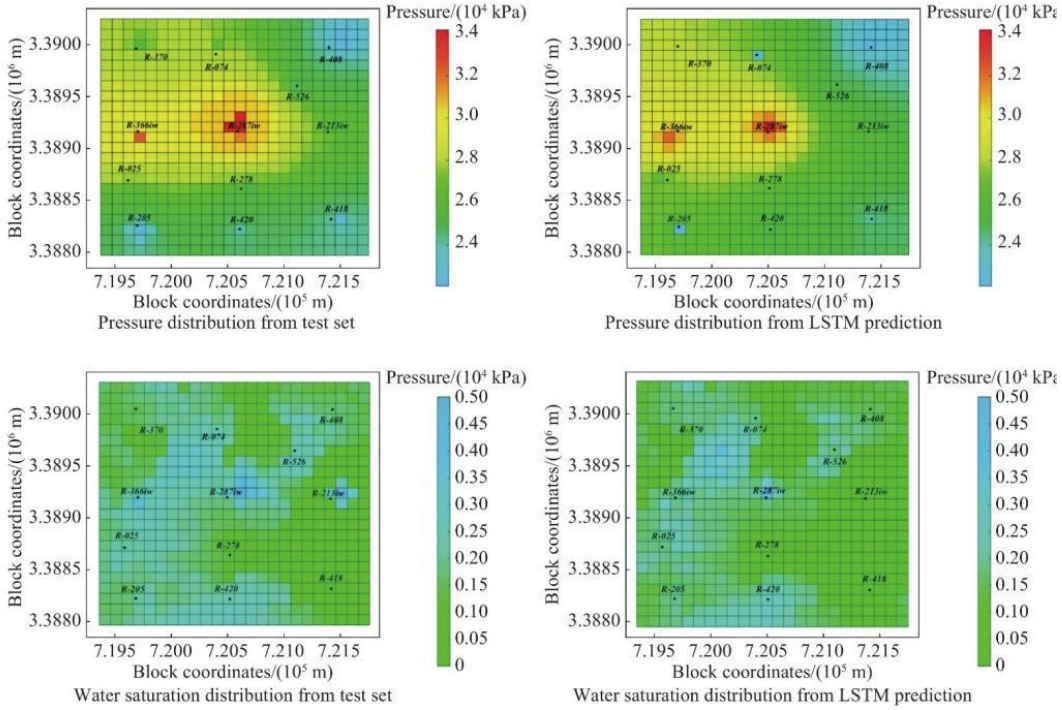

Figure 6: Prediction effect of remaining oil distribution based on LSTM neural network

4.3. China's oil industry development needs to apply big data

At present, there are several main features: First, the big data analysis platform of large oil companies is being or basically completed. Domestic and foreign energy companies are trying to use big data analysis technology to carry out oil and gas exploration and development business, and have achieved good implementation results. By establishing a big data analysis platform, connecting people, processes and technologies, the value can be maximized; second, small wins big. Through big data, the changes of various related factors can be discovered and their results can be predicted, thereby improving business foresight, aiming to solve specific business problems and realize automatic optimization of business; third, the introduction of intelligent algorithms. By introducing intelligent algorithms such as machine learning and deep learning, the accuracy and real-time performance of big data analysis are further enhanced.

With the trend of increasingly extensive and in-depth application of big data technology, the oil industry should join hands with Internet technology to clarify the application needs of various fields and actively promote the promotion and application of big data technology. The specific suggestions are as follows: 1. Establish a unified oil engineering big data platform with data interoperability and information sharing. Data is the basis for the intelligent transformation of oil enterprises. The establishment of a unified big data platform requires the joint participation of oil field branches, oil field service companies, scientific research units, etc. Breaking the situation of data silos, dispersion, and mutual isolation, realizing information data sharing between different professions and departments, and standardizing data collection, transmission, storage, conversion, integration, and application are the only way to strengthen data sharing in all aspects of oil engineering. Standardize data collection, transmission, and storage before data fusion to improve data consistency and reliability. 2. Fusion of forward-looking technologies. In

ISSN 2616-5767 Vol.8, Issue 1: 24-31, DOI: 10.25236/AJETS.2025.080104

view of the problem that the performance of traditional big data optimization algorithms drops sharply when processing high-dimensional oil engineering big data, combined with the characteristics of oil engineering big data, strengthen the basic research of intelligent algorithms such as machine learning and deep learning, carry out forward-looking technology research such as drilling wellbore digital twins, drilling and completion robots, well site virtual reality/augmented reality, and cognitive computing, improve the speed and accuracy of oil engineering big data analysis, and lay a technical foundation for the next generation of intelligent development of oil engineering. 3) Establish a complete data security defense system. The oil industry involves a large amount of sensitive data. How to ensure data security and privacy protection is an important challenge. While developing big data applications, it is also necessary to prevent data leakage to prevent the outflow of important data from leading to other data security threats and network attacks, and destroying the integrity and authenticity of the data.

5. Conclusion

The three Asian countries of China, India, and Pakistan have all made certain progress and achievements in the application of big data technology in the petroleum field. China is at the forefront of big data applications with its huge market size and advanced technical strength; India, with technological innovation and policy support, continues to promote the popularization of big data technology in the oil industry; and Pakistan is also actively catching up with international trends and increasing Big technology introduction and cooperation efforts. In the future, with the continuous development and improvement of big data technology, the above three countries are expected to achieve more significant results and breakthroughs in the field of big data applications in the petroleum field. At the same time, more countries are expected to join the ranks of big data technology applications.

References

[1] Ahmad, K., Maabreh, M., Ghaly, M., Khan, K., Qadir, J., & Al-Fuqaha, A. Developing future human-centered smart cities: Critical analysis of smart city security, Data management, and Ethical challenges[J]. Computer Science Review, 2022,43, 100452.

[2] Z. Latif, M. Z. Tunio, Z. H. Pathan, Z. Jianqiu, L. Ximei and S. K. Sadozai, "A review of policies concerning development of big data industry in Pakistan: Subtitle: Development of big data industry in Pakistan," 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, Pakistan, 2018, pp. 1-5.

[3] Verma, P., & Soni, P. Role of Big Data Analytics in Indian Oil and Gas Industry: Challenges and Opportunities[J]. International Journal of Emerging Technology and Advanced Engineering, 2021,11(1), 19-25.

[4] Ashootosh Chand, Abhishek Verma & Debankur Gosh. Oil and Gas Digital Transformation using edge computing: PWC Journal of Oil and Gas,2023.

[5] Hassani H, Silva E S. Big Data: a big opportunity for the petroleum and petrochemical industry[J]. OPEC Energy Rev, 2018, 42(1): 74.

[6] SONG Hong-qing, DU Shu-yi, ZHOU Yuan-chun, WANG Yu-he, WANG Jiu-long. Big data intelligent platform and application analysis for oil and gas Big data intelligent platform and application analysis for oil and gas resource development[J]. Chinese Journal of Engineering, 2021, 43(2): 179-192.

[7] YANG Yong.Application progress of big data & AI technologies in exploration and development of Shengli Oilfield[J]. Oil and Gas Geology and Recovery, 2022, 29(1):1-10.

[8] Wu X, Zhong G .Generating a relative geologic time volume by 3D graph-cut phase unwrapping method with horizon and unconformity constraints[J]. Geophysics, 2012, 77(4):21. doi:10.1190/geo2 011-0351.1.

[9] LOU Y, ZHANG B, LIN T, et al. Seismic horizon picking by inte- grating reflector dip and instantaneous phase attributes[J].Geo- physics,2020,85(2):O37-O45.

[10] Figueiredo A M, Silva F B, Silva P M, et al. A Clustering-based Approach to Map 3D Seismic Horizons[C]//14th International Congress of the Brazilian Geophysical Society & EXPOGEF, Rio de Janeiro, Brazil, 3-6 August 2015.

[11] Stark T J .Relative geologic time (age) volumes-Relating every seismic sample to a geologically reasonable horizon[J].Leading Edge 2004, 23(9):928.

[12] Zaremba W, Sutskever I, Vinyals O. Recurrent neural network regularization. arXiv preprint (2015-02-19) [2020-07-07]. https://arxiv.org/abs/1409.2329.

[13] Zhang Q T, Wei C J, Wang Y H, et al. Potential for prediction of water saturation distribution in reservoirs utilising machine learning methods[J]. Energies, 2019, 12(19): 3597.