

Data Science Applications in Oil & Gas Industry

The Oil and Gas industry can benefit greatly from the latest advancements in data science. Digital transformation represents their winning card while exploiting new oil and gas assets and getting advantage of Hyperconnectivity with cloud and real-time intelligence to streamline the performance of their existing extraction, refining and distribution global operations amid tough market conditions and timely address environmental priorities, such as green gas emission reduction to comply with Paris Climate Change Agreement in 2030.

In this article, I will conduct an overall review on the applications of Data Science in the oil and gas industry, and illustrate how Data Science plays a key role in simplify oil and gas operations during the three phases of oil life-cycle process:

- Upstream (exploration, drilling, reservoir, production).
- Midstream (transportation).
- Downstream (refining, delivery).

Big Data Analytics Use Cases In The Oil & Gas Industry



Upstream Phase

The upstream sector is the first phase of oil and gas extraction process, it contains the exploration and drilling, reservoir and production operations. The application of Data science has become prominent as the amount of generated and recorded data has dramatically increased during the upstream phase. The improvements in seismic acquisitions devices, channel counting, fluid front monitoring geophones, LWD, MWD and wireline tools have provided a huge amount of data to be processed and analyzed. I will summarize how data science contributes to each upstream operation and presents an informative insight from the vast amount of available data in this phase.

I. Seismic Data

The processing and interpreting of seismic data require sophisticated processing computers with powerful visualizations capabilities. With the recent improvements in seismic devices, the amount of generated data has boosted significantly. The detailed interpretation of these new datasets needs to go beyond the conventional methods. One of the most important applications of Data Science in the oil and gas industry is analyzing the seismic data [\[1\]](#). Machine learning tools can reveal the relationship between the recorded data more efficiently, specifically when dealing with huge datasets. In research conducted by Roden [\[2\]](#), the author

incorporated principal component analysis (PCA) with self-organizing maps (SOM) to carry out multi-component seismic analysis. In his research, the analysis contains five stages as follow:

- 1- The geological issue was clearly defined.
- 2- PCA was run to identify the key attributes related to the defined problem.
- 3- SOM was run by employing machine learning tools to train a prediction model.
- 4- The outcomes of SOM analysis were further analyzed by 2D maps to identify the important geological features.
- 5- Sensitivity analysis was conducted to refine the results by considering various attributes and different training model scenarios.

II. Drilling Optimization

One method to increase the drilling efficiency is to create machine learning models that predict possible equipment failures. The equipment has sensors around the drilling rig for collecting data during drilling operations, which is transmitted in real-time-to-Real-Time Operation Center (RTOC) in the head quarter offices. This data can be used to apply machine learning algorithms and predict any deviation from the performance benchmark. In a study by Maidla et al. [3] the drilling performance was improved by applying Data analytics and including drilling and formation parameters. In their study, the data from morning report, electronic drilling recorder (EDR), and cross-plots of weight on bit (WOB) and differential pressure were used to optimize the drilling performance.

III. Reservoir Engineering

New generation of reservoir simulation technique incorporates the artificial intelligence and data mining technologies with the Closed-Loop Reservoir Management (CLRM) and Integrated Asset Modeling (IAM). The result will be an innovative information-oriented reservoir modeling approach, which can improve the modeling by predicting the affective parameters which theory-based equations of state cannot capture. Bello et al. [4] used the data of reservoir characterization to develop a reservoir management application based on utilizing Data Science tools, including data visualization, downhole data preprocessing, create machine learning model and do the training, development, and model validation. They used the Apache Spark machine learning tool to conduct their project and they deployed the developed model to the cloud to facilitate the user/system interactions.

IV. Production Engineering

Seemann et al. [5] from Saudi Aramco developed a smart forecast and flow method to conduct automated decline analysis of oil production. Their goal was to identify the underlying pattern in production data and to forecast the production performance.

Midstream Phase

Oil and gas's transportation represent a big challenge to the oil producers around the world. That challenge demonstrates in assure carry their products in a safe and secure way without

any risk in order to guarantee environment protection and avoid any possible liability from the government jurisdiction agencies. Therefore, pipeline's companies using spread sensors along the pipelines and GPS devices on tankers to ensure secure energy product transportation. With large-scale data monitoring automation system, organizations can get continuous monitoring on tanker and pipeline sensor's data to detect any leakage, spills or malfunctions such as stress corrosion, fatigue fractures, earthquake displacement, etc.

Downstream Phase

Oil and gas companies may use data science to minimize downtimes, refining equipment maintenance costs and boosting asset management. A comparison of the past and present operational data of the equipment analyzes its performance. The performance estimate is further adjusted according to the device's end-of-life criteria and failure situations. In a recent project by Repsol SA, Data Science is utilized to conduct management optimization for one of the company's refineries in Spain. For this project, Google Cloud would provide Repsol with data analytics products and consultation as well as Google Cloud machine learning services [6]

Conclusion

The oil & gas industry is one of the largest industries in the global economy. The population is rising across the world and the demand for oil and gas is growing tremendously. Data Science has gained interest and attention of big oil companies around the world, and they started adapting Data Science techniques in their moving towards the Digital Transformation, AI and machine learning era.

References

1. A. Alfaleh, S. Aramco, Y. Wang, A. Texas, B. Yan "Topological Data Analysis to Solve Big Data Problem in Reservoir Engineering: Application to Inverted 4D Seismic Data (2015)"
2. R. Roden, G. Insights "Seismic Interpretation in the Age of Big Data (2016), pp. 4911-4915"
3. E. Maidla, W. Maidla, J. Rigg, M. Crumrine, P. Wolf-zoellner "Drilling analysis using big data has been misused and abused IADC/SPE Drill. Conf. Exhibit., Fort Worth (2018)"
4. O. Bello, D. Yang, S. Lazarus, X.S. Wang, T. Denney, B.H. Incorporated "Next Generation Downhole Big Data Platform for Dynamic Data-driven Well and Reservoir Management (2017)"
5. D. Seemann, M.W. SPE, S. Hasan, S. Aramco "SPE 167482 Improving Reservoir Management through Big Data Technologies (2013), pp. 28-30"
6. R. Brelsford "Repsol launches big data, AI project at Tarragona refinery Oil Gas J., 116 (2018)"