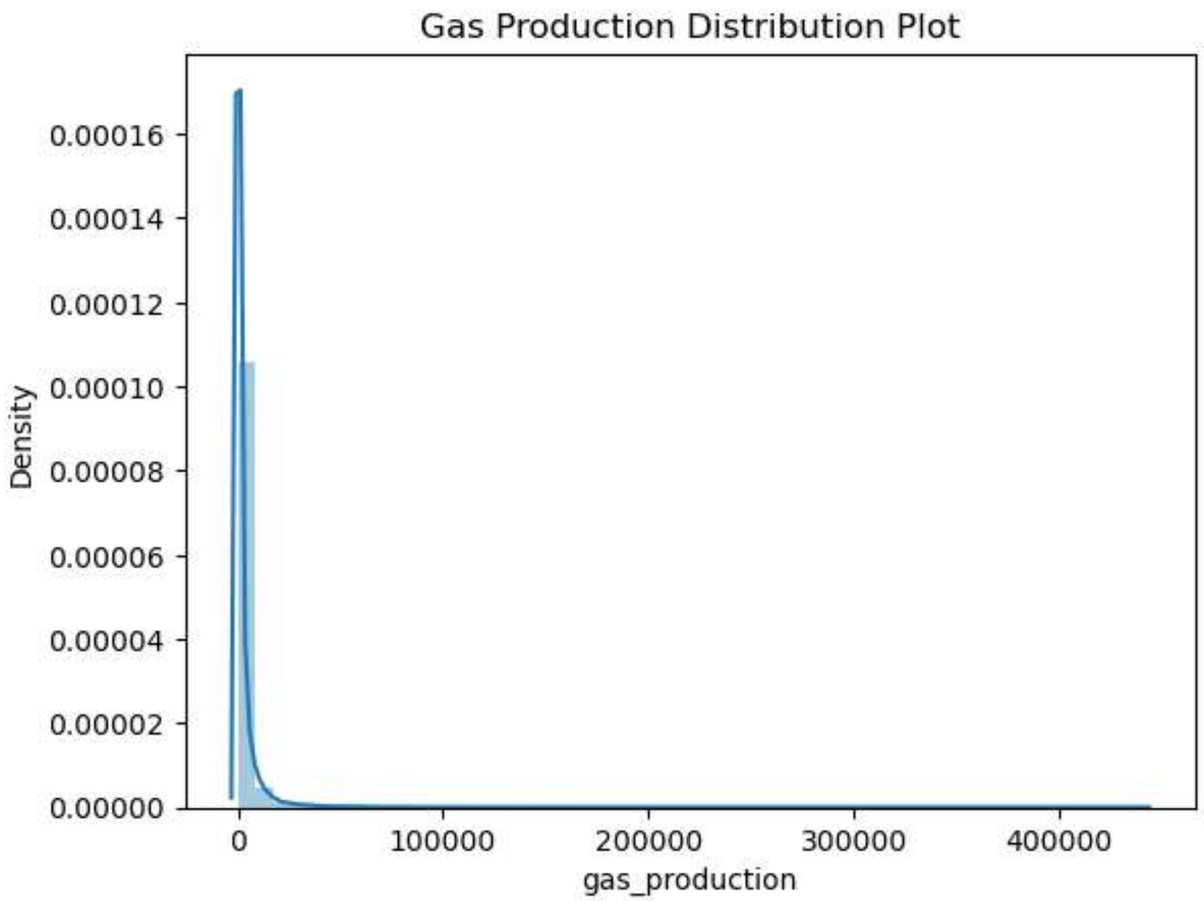
Oil and Gas AI-ML Case Study

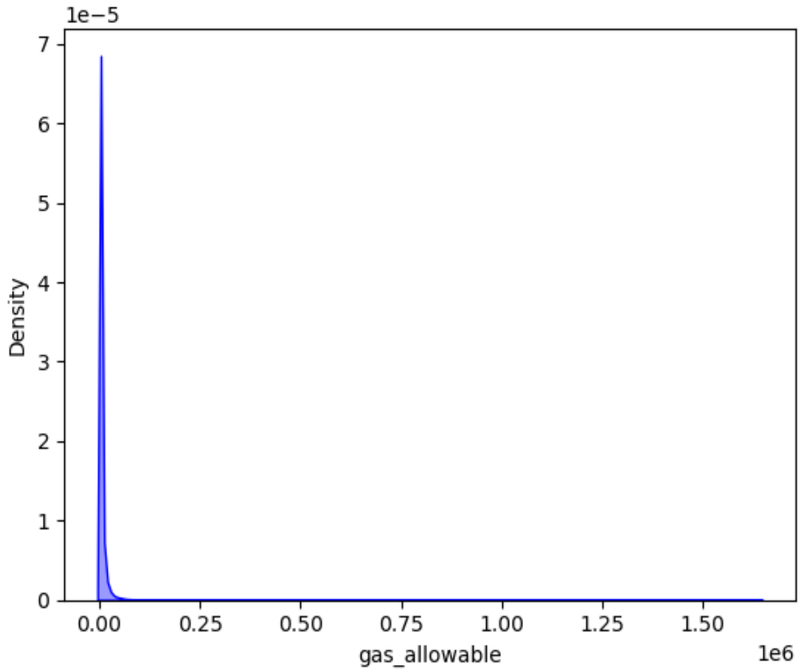
The oil and gas industry is a complex and challenging field that involves the exploration, production, and processing of hydrocarbons. Auto-ML has the potential to play a significant role in this industry by helping to improve operational efficiency, reduce costs, and mitigate risk. Auto-ML algorithms can be trained on vast amounts of historical data to predict and optimize production, identify equipment failures, and minimize downtime. This can lead to significant cost savings and improved safety. Auto-ML can also be used to improve reservoir modeling, by better predicting the geological and physical properties of oil and gas reservoirs. This can lead to more accurate predictions of reservoir performance and enable better decision making.

The aim of this study is to predict Gas production and identify factors that optimize Gas production using Auto-ML & AI.

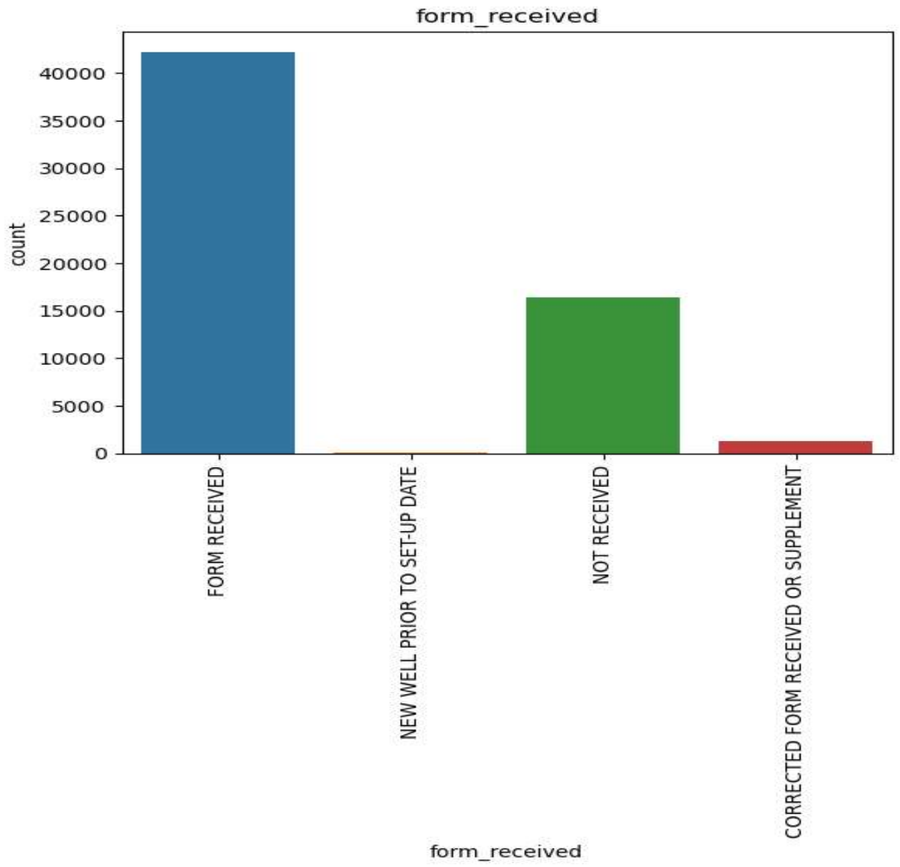


- The business feature used here is gas_production.
- The distribution plot shows that most of the gas_production is from 0 - 100000.

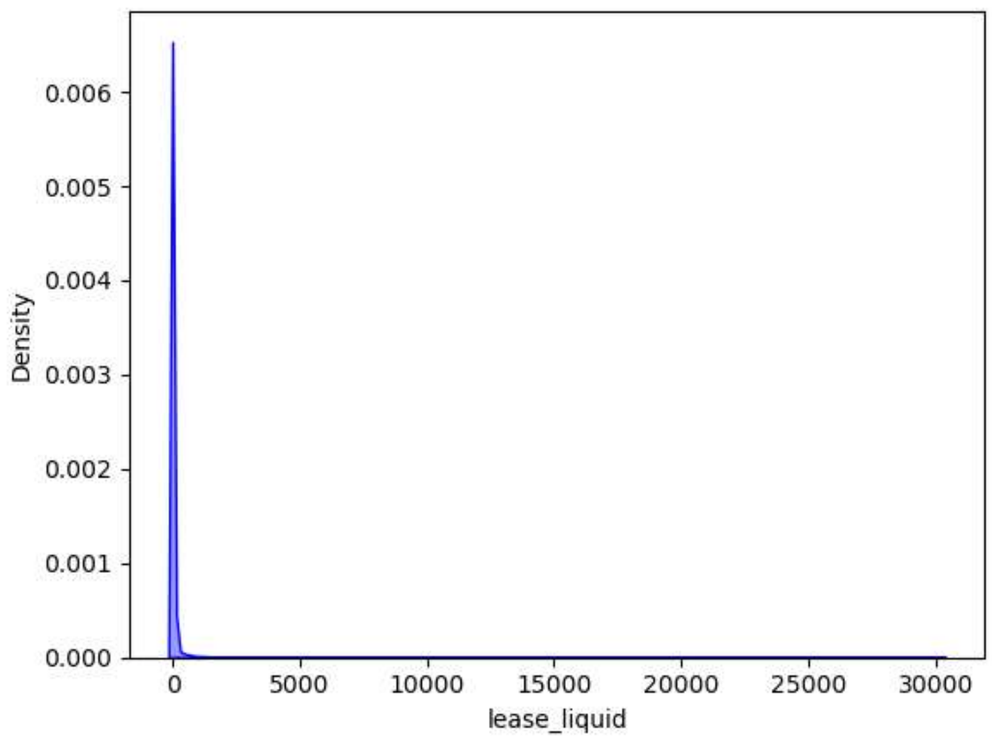
Features Responsible



- **gas_allowable** = Gas allowable refers to the amount of natural gas that is permitted to be produced from a particular well or field over a specified period of time.



- **form_received** = The form recieved or not to start a new well for gas production



- **lease_liquid** = Lease liquids refer to the natural gas liquids (NGLs) and oil that are produced along with natural gas from a well or field. These liquids can include ethane, propane, butane, pentane, and heavier hydrocarbons.

Auto-ML Methodology Results

Algorithms	Test Accuracy (25 percentile)	Test Accuracy (50 percentile)	Test Accuracy (75 percentile)	Test Accuracy (90 percentile)
Lasso	72.3	72.7	72.8	72.8
Random Forest	86.8	86.7	87.5	87.3
XGBoost	82.8	78.5	80.2	80.1
RNN	76.2	75.3	71.8	74.1
ANN	71.3	73.5	65.7	63.9
Total Features	17	34	51	61
Avg. Accuracy	77.88	77.34	75.54	75.64

- **Based on our observation from the standard ML algorithms, 25th percentile has the best average accuracy.**
- **Random Forest was the best performing algorithm with 87.5% accuracy in 75 percentile.**

Conclusion

Auto-ML algorithms can be trained on historical data to predict when equipment is likely to fail or require maintenance. This can help reduce downtime and improve operational efficiency. Auto-ML can be used to optimize oil and gas production by analyzing large amounts of data from sensors and other sources to identify patterns and make predictions. The dataset has 60,000 records with 17 Categorical Features and 26 Numerical Features.

For regression, models were created with algorithms using Auto-ML techniques like Lasso, Random forest, XGBoost, Recurrent Neural Network and Artificial Neural Network. With these models, performance measurement values were obtained for feature sets of 17, 34, 51 and 61. The Auto-ML algorithms were able to predict thegas production with an average accuracy between 75% – 78% and helped to identify factors that optimize the gas production. The major factors include Gas_allowable, Form_Received_Not_Received and Lease liquid. The Random forest with 87.5 % accuracy in 75th percentile where tree showed a threshold of gas allowable >= 175636.5 units which leads to highest gas production.

Overall, Auto-ML can help solve a range of problems in the oil and gas industry, from improving efficiency and reducing costs to enhancing safety and mitigating environmental impacts.