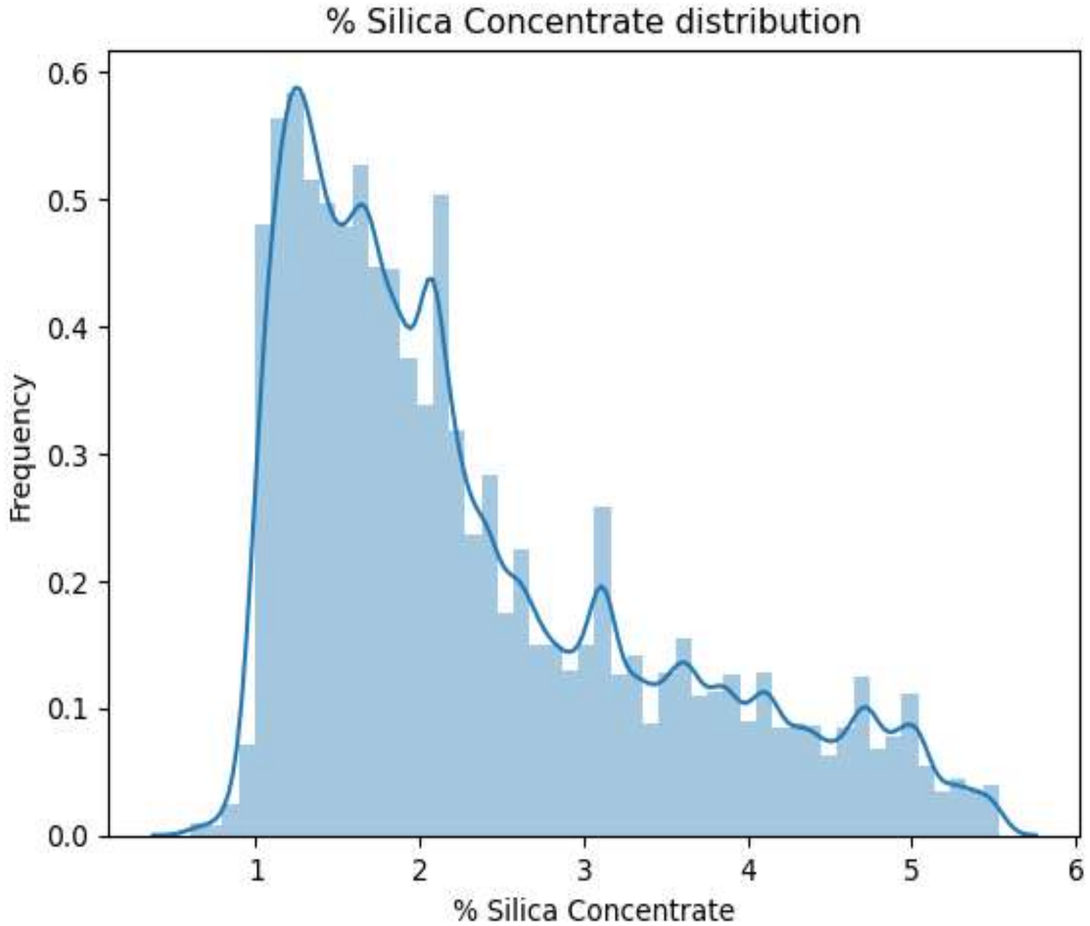
Mining AI-ML Case Study

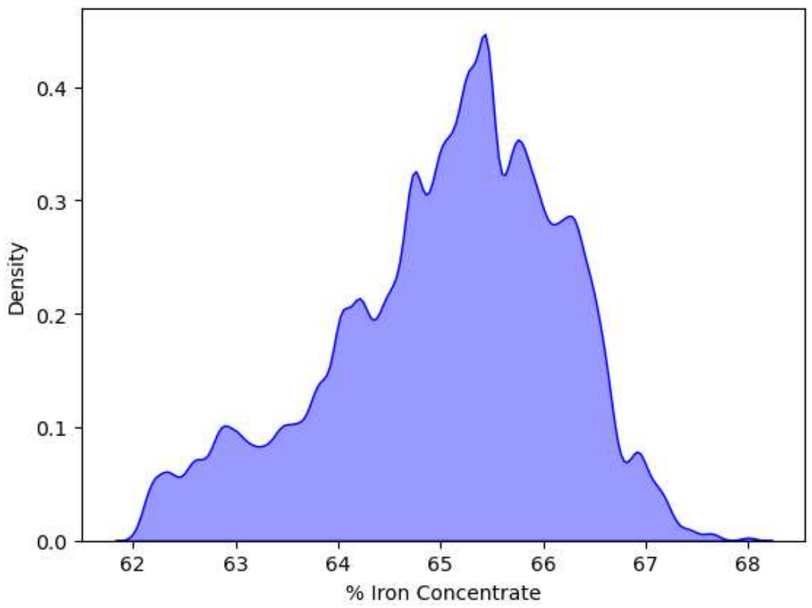
The main goal is to use this data to predict how much impurity is in the ore concentrate. As this impurity is measured every hour, if we can predict how much silica (impurity) is in the ore concentrate, we can help the engineers, giving them early information to take actions. Hence, they will be able to take corrective actions in advance (reduce impurity, if it is the case) and also help the environment (reducing the amount of ore that goes to tailings as you reduce silica in the ore concentrate).

The dataset contains 737,453 observations. The aim of this study is to predict the impurity of the iron ore concentrate and identify the features determining impurity using Auto-ML & AI.

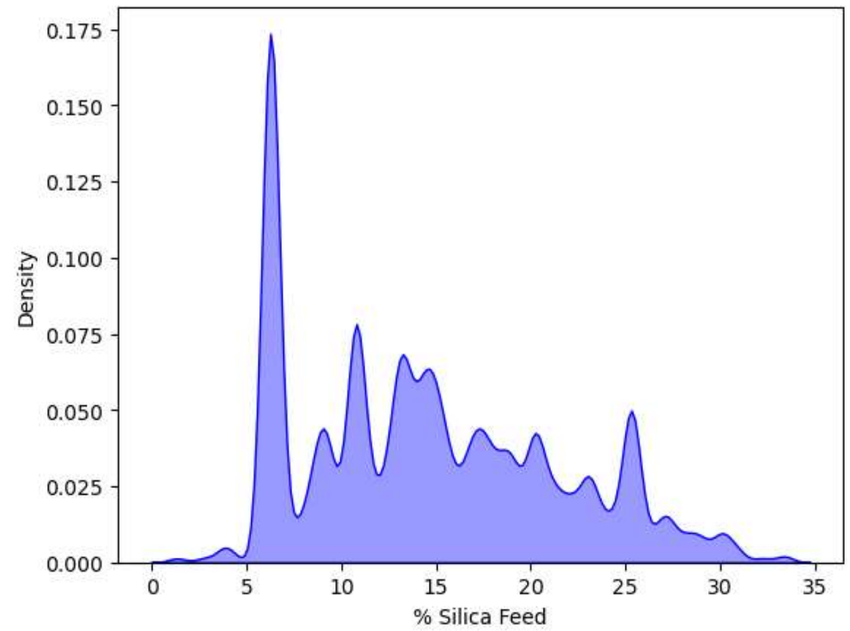


- The business feature used here is % Silica Concentrate.
- % of silica which represents how much iron is presented in the end of the flotation process (0-100%, lab measurement)
- The distribution plot shows that most of the % Silica Concentrate present in the ore is from 1 – 3 %.

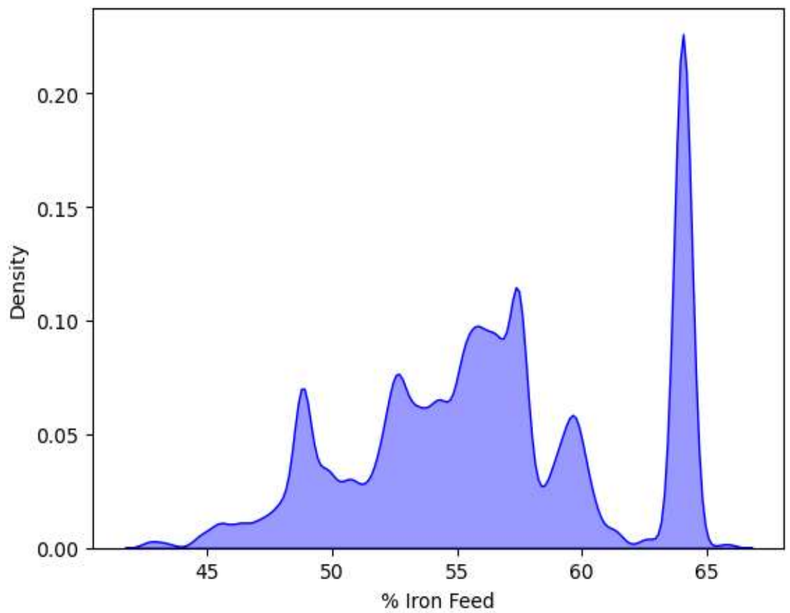
Features Responsible : Auto-ML



- **% Iron Concentrate = % of Iron which represents how much iron is presented in the end of the flotation process (0-100%, lab measurement)**



- **% Silica Feed** = % of silica (impurity) that comes from the iron ore that is being fed into the flotation cells



- **% Iron Feed** = % of Iron that comes from the iron ore that is being fed into the flotation cells

Auto-ML Methodology Results

Algorithms	Test Accuracy (25 percentile)	Test Accuracy (50 percentile)	Test Accuracy (75 percentile)	Test Accuracy (90 percentile)
Lasso	66.5	66.8	68.2	68.2
Random Forest	99.6	99.9	99.9	99.9
XGBoost	95.4	95.5	95.4	95.6
MLP	71.6	73.4	66.1	71.8
RNN	66.2	46.2	67.3	7.5
Total Features	6	12	18	22
Avg. Accuracy	79.86	76.36	79.38	68.6

- **Based on our observation from the standard ML algorithms, 25th percentile has the best average accuracy.**
- **Random Forest was the best performing algorithm with 99.9% accuracy in 50, 75 and 90 percentile.**

Conclusion

Auto-ML can be a powerful tool for solving problems in the mining industry. Auto-ML algorithms can analyze geological data and identify patterns that may indicate the presence of valuable minerals. This can help mining companies to target their exploration efforts more effectively and reduce the time and cost involved in finding new mineral deposits. Auto-ML models can analyze data from sensors and other monitoring equipment to identify potential equipment failures before they occur which help in schedule maintenance proactively, reducing downtime and increasing the lifespan of equipment. The dataset has 737,453 records with 1 Categorical Features and 25 Numerical Features.

For regression, models were created with algorithms using Auto-ML techniques like Lasso, Random forest, XGBoost, Multilayer Perceptron and Recurrent Neural Network. With these models, performance measurement values were obtained for feature sets of 6, 12, 18 and 22. The Auto-ML algorithms were able to predict the impurity of the iron ore with an average accuracy between 67% – 80% and helped to identify factors that determine the imnpurity in iron ore.The major factors that determine impurity in iron ore is % Iron Concentrate, % Silica Feed and % Iron Feed. The Random forest with 99.9 % accuracy in 90th percentile where tree showed a threshold of Iron concentrate <= 63.54% which leads to highest Silica Concentrate.

Overall, Auto-ML has the potential to help mining companies reduce costs, improve efficiency, and increase the safety of their operations.