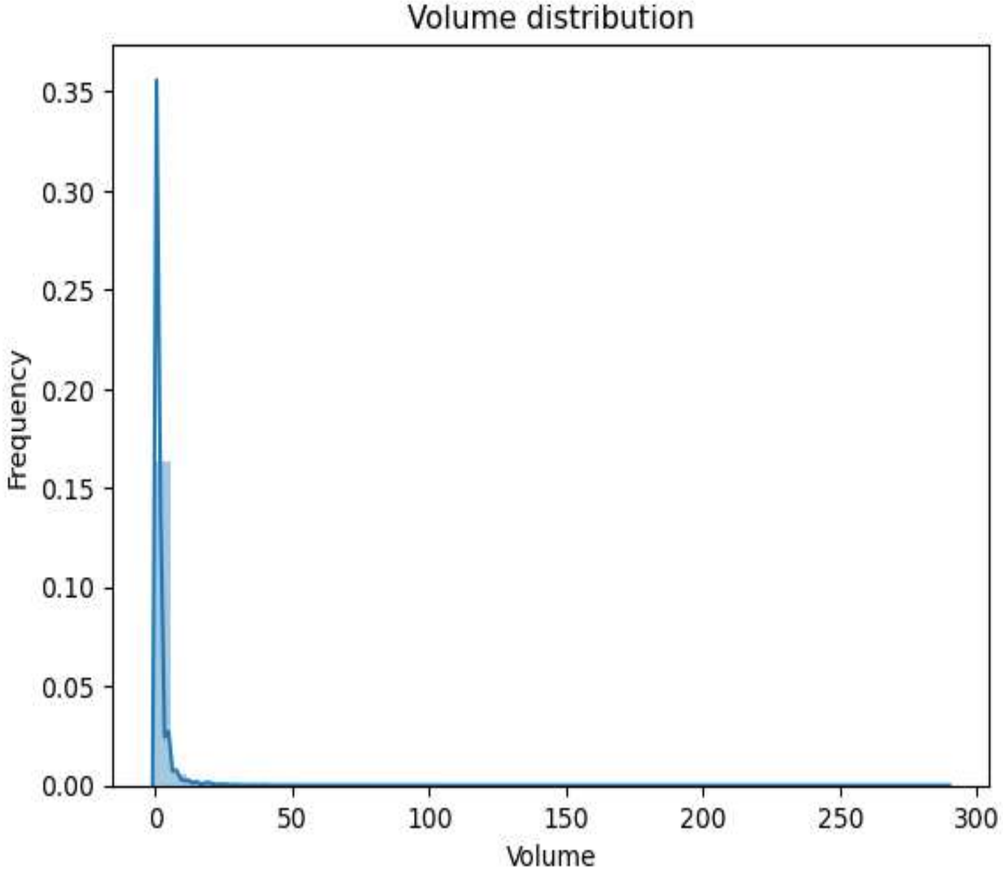
Cargo AI-ML Case Study

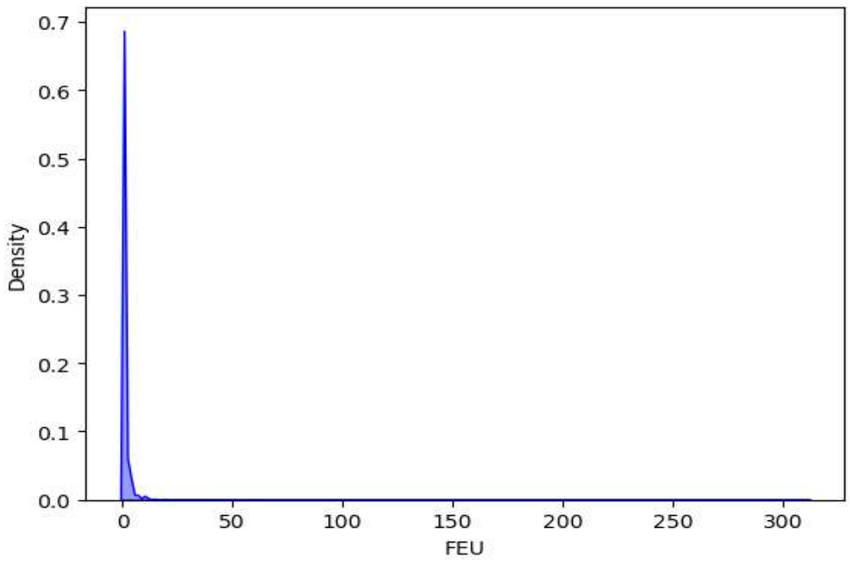
The dataset contains information about the shipment of various goods via ship. Shipping goods via ship, also known as maritime transport, is one of the most commonly used modes of transportation for international trade. It involves the transportation of goods and commodities via cargo ships, which are large vessels designed for carrying goods in bulk or in containers. The shipping cargo industry generates vast amounts of data that can be analyzed using machine learning techniques to extract valuable insights and optimize operations. Auto-ML algorithms can be trained on historical data to identify patterns and trends, and to make predictions about future events.

The aim of this study is to predict the volume of the cargo and find the factors which determine volume using ML & AI.

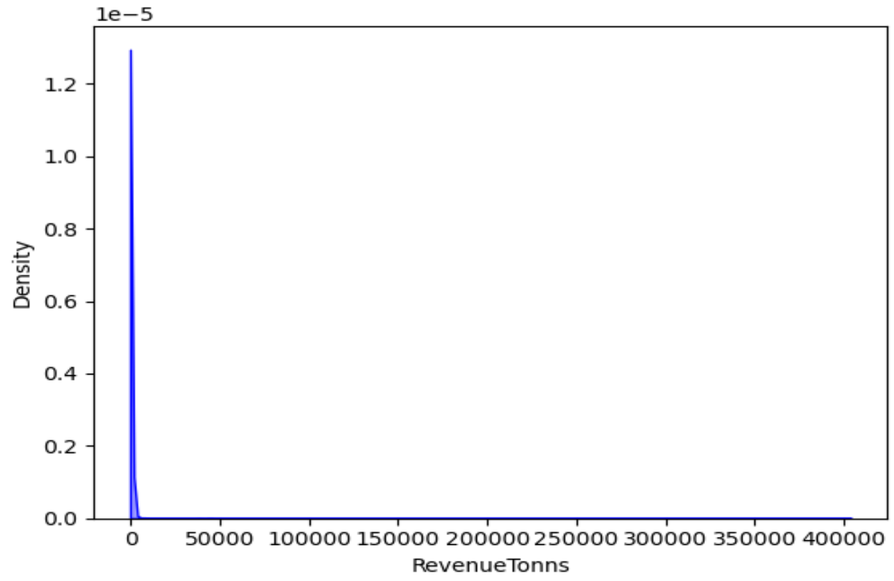


- The business feature used here is Volume.
- Volume is the space taken by goods in a cargo being shipped
- The distribution plot shows that most of the volume is below 50 units.

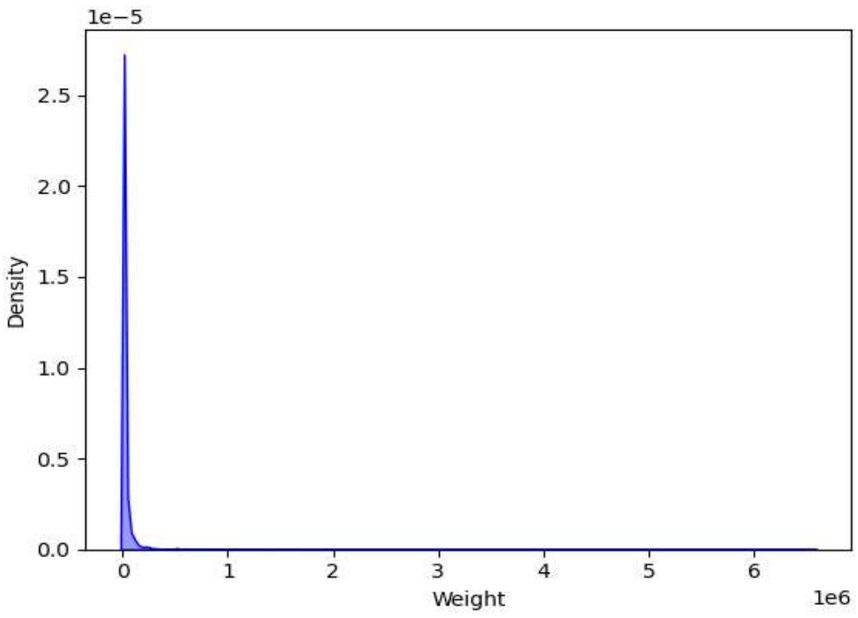
Features Responsible



- **FEU =** Forty-foot Equivalent Unit
- The forty-foot equivalent unit (FEU) is a unit of measurement used to describe the volume of cargo capacity in intermodal containers. It refers to the size of a standard shipping container, which is typically 40 feet long, 8 feet wide, and 8 feet 6 inches high.



- **RevenueTonnes =** Revenue Tons (RT) is a unit of measurement used in the shipping industry to determine the volume of cargo carried on a vessel.
- It refers to the total weight of cargo that can be carried on a ship, taking into account both the weight of the cargo and the space it occupies.



- **Weight =** Weight is the measure of the mass or heaviness of the cargo. However, weight can indirectly impact the volume of the shipping cargo through weight limitations and capacity constraints.

Auto-ML Methodology Results

Algorithms	Test Accuracy (25 percentile)	Test Accuracy (50 percentile)	Test Accuracy (75 percentile)	Test Accuracy (90 percentile)
Lasso	87.3	87.3	87.4	87.6
Random Forest	98.6	98.6	98.7	98.6
XGBoost	98.6	98.5	98.4	98.4
MLP	89.17	96.2	96.8	89.4
RNN	96.2	89.17	52.09	33.13
Total Features	7	14	21	25
Avg. Accuracy	93.974	93.954	86.678	81.426

- **Based on our observation from the Auto-ML algorithms, 25th percentile has the best average accuracy.**
- **Random Forest was the best performing algorithm with 98.7% accuracy in 75 percentile.**

Conclusion

AI can monitor data from various sources such as weather, sea conditions, and vessel traffic to provide a real-time data analysis, decision support, and automated alerts to help ship operators navigate safely and efficiently. Auto-ML models can be used to analyze data and make predictions about potential hazards or risks, such as storms, collisions, or equipment failures. The dataset has 278,531 records with 21 Categorical Features and 8 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 7, 14, 21 and 25. The Auto-ML algorithms were able to predict the volume of the cargo with an average accuracy between 80% – 94% and helped to identify factors that determine the volume. The major factors include Forty-foot Equivalent Unit (FEU), RevenueTonnes and Weight. The Random forest with 98.7 % accuracy in 75th percentile where tree showed a threshold of Forty-foot Equivalent Unit >= 9.44 units and Weight >= 2143522 units which leads to highest volume.

Overall, AI can be used to optimize cargo loading and unloading, by analyzing data on cargo weight and volume, vessel stability, and port infrastructure.