**# Industrial-Copper-Modeling-Project**

**## Introduction**

This project aims to develop two machine learning models for the copper industry to address the challenges of predicting selling price and lead classification. Manual predictions can be time-consuming and may not result in optimal pricing decisions or accurately capture leads. The models will utilize advanced techniques such as data normalization, outlier detection and handling, handling data in the wrong format, identifying the distribution of features, and leveraging tree-based models, specifically the decision tree algorithm, to predict the selling price and leads accurately.

**## Regression model details**

The copper industry deals with less complex data related to sales and pricing. However, this data may suffer from issues such as skewness and noisy data, which can affect the accuracy of manual predictions. Dealing with these challenges manually can be time-consuming and may not result in optimal pricing decisions. A machine learning regression model can address these issues by utilizing advanced techniques such as data normalization, outlier detection and handling, handling data in wrong format, identifying the distribution of features, and leveraging tree-based models, specifically the decision tree algorithm.

**## Classification model details**

Another area where the copper industry faces challenges is in capturing the leads. A lead classification model is a system for evaluating and classifying leads based on how likely they are to become a customer. You can use the STATUS variable with WON being considered as Success and LOST being considered as Failure and remove data points other than WON, LOST STATUS values.

**## Solution**

The solution includes the following steps:

Exploring skewness and outliers in the dataset.

Transforming the data into a suitable format and performing any necessary cleaning and pre-processing steps.

Developing a machine learning regression model which predicts the continuous variable 'Selling\_Price' using the decision tree regressor.

Developing a machine learning classification model which predicts the Status: WON or LOST using the decision tree classifier.

Creating a Streamlit page where you can insert each column value and get the Selling\_Price predicted value or Status (Won/Lost).

**## Requirements**

This project requires the following libraries to be installed:

NumPy

Pandas

Scikit-learn

Streamlit

**## Getting Started**

Clone the repository.

Install the required libraries.

Run the Streamlit app using the command: streamlit run app.py.

Enter the values for each column to get the Selling\_Price predicted value or Status (Won/Lost).