**INTRODUCTION**

This project seeks to create machine learning model tailored for the copper industry, with the primary goal of addressing the difficulties associated with predicting selling prices and classifying leads. Manual predictions can be both time-consuming and prone to inaccuracies, making it challenging to make optimal pricing decisions or correctly identify leads. The proposed models will incorporate advanced techniques, including data normalization, outlier detection and treatment, addressing data format issues, exploring feature distributions, and harnessing tree-based models to achieve precise predictions for both selling prices and lead classification.

**Prerequisites**

1. **Python** -- Programming Language
2. **pandas** -- Data Preprocessing
3. **Matplotlib and Seaborn** --Data Visualisation
4. **Streamlit** -- User Interface
5. **scikit-learn** -- Machine Learning library for the Python programming language

**Project Workflow**

The following is a fundamental outline of the project:

* This analysis aims to investigate the presence of skewness and outliers within the dataset.
* The data will be converted into a format that is appropriate for analysis, and any required cleaning and pre-processing procedures will be carried out.
* The objective of this study is to construct a machine learning regression model that utilizes the decision tree regressor to accurately forecast the continuous variable 'Selling\_Price'.
* The objective of this study is to construct a machine learning classification model using the decision tree classifier to accurately predict the outcome of a given task, namely whether it will result in a "WON" or "LOST" status.
* The objective is to develop a Streamlit webpage that enables users to input values for each column and get the expected Selling\_Price value or determine the Status (Won/Lost).