**Stock Market Prediction on Tesla**

Our project is focused on stock price prediction, and our team consists of

Abhishek Thamatam

Pranay Kumar Pokala,

Sneha Reddy Madhireddy

Akhileshwar Rao Emmanani

Shashank Koya

The goal of our project is to develop a machine-learning model that accurately predicts the future prices of a given stock based on historical data.

To achieve this goal, we aim to answer several research questions, including:

* How can we best preprocess and clean our dataset to ensure accurate predictions?
* What are the most important features that impact stock prices, and how can we incorporate them into our model?
* Which machine learning algorithms and techniques are most effective for stock price prediction?

We plan to use historical stock price data from Yahoo Finance as our primary dataset. Specifically, we will focus on a single stock Tesla and collect daily closing prices for several years. Additionally, we may explore incorporating other datasets such as news articles or social media sentiment to further refine our model.

The total size of our dataset will depend on the number of years we choose to collect data for, but we expect to have several thousand rows of data in total. Each row will represent values which contain the opening and closing stock, day, volume, etc.

Potential features for our research problem include company financial data, market trends, stock market indexes, and external factors that can impact stock prices. The potential features we have identified include stock volume, and opening and closing prices. We will need to carefully evaluate and select the most impactful features to include in our model and may use techniques such as feature engineering or PCA to reduce the dimensionality of our dataset.

Our target variable will be the stock price. It is a continuous variable, and our goal is to predict the future prices of the stocks accurately.

In this, we are going to use supervised algorithms where our data have features which are labelled and the number of columns is 7 columns and the years we are going to take is from 2010 to 2022.