**AI and ML Engineer Task Report: Week 1 & 2**

### Week 1: Data Preparation and Initial Model Training

**Task 1: Dataset Preparation** For the stock market prediction task focusing on Apple and Samsung, I created a comprehensive dataset that includes the following key features:

* **Price**: Historical stock prices for Apple and Samsung.
* **Value for Money**: A calculated metric based on sales price and features.
* **Total Profits**: Cumulative profits generated by each company.
* **Total Number of People in Use**: The total number of users engaged with the company's services.
* **Total Sales Since Launch**: Aggregated sales figures since each product's launch.
* **Best Seller in Country**: Identification of the top-selling country for each company based on stock sales.
* **Best in Tech**: A comparative metric indicating which company is leading in technology sales.
* **Stock Value**: The current market value of stocks.
* **Total Number of Average Stocks Bought**: The total average number of stocks bought up until today.

**Task 2: Data Cleaning and Model Training**

* **Data Cleaning**: I ensured the dataset was free from duplicates, which could skew the model’s predictions. Any duplicates found were promptly removed.
* **Model Training**: I applied various machine learning algorithms to predict stock trends, starting with a Decision Tree model. I evaluated its performance using accuracy metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared score.

**Model Performance**:

* **Decision Tree Model**: The model performed well, with a Mean Absolute Error of 0.43 units, indicating high prediction accuracy. The R-squared score of 98.4% demonstrated that the model effectively captured the variance in the dataset, particularly for predicting the cheaper price if the same product were made with all raw materials available in India.

### Week 2: Model Visualization and Advanced Predictions

**Task 1: Data Visualization** Using the dataset from Week 1, I created visualization charts to better understand the model’s performance. I employed various charts:

* **Pie Charts**: To display the proportion of sales across different countries.
* **Line Plots**: For visualizing the trends in stock prices over time.
* **Bar Plots**: To compare the total profits and sales between Apple and Samsung.

**Task 2: Advanced Predictions** I extended the analysis by predicting the percentage of success if the same product was made with all raw materials available in India at a cheaper price. Additionally, I explored the product’s global efficiency based on daily stock values.

**Task 3: Model Accuracy Assessment**

* **Week 2 Model Performance**: The model showed a significant deviation in predictions during Week 2, with a Mean Absolute Error of 183.56 units, indicating that predictions were less reliable. Despite a high R-squared score of 99.3%, the large errors suggested inconsistencies, particularly when simulating the cheaper product scenario because of the data values were random.

### Summary of Learnings

Over the course of these two weeks, I gained valuable insights into the importance of data quality and the impact of feature engineering on model accuracy. I learned how to effectively use Decision Tree algorithms and the importance of parameter tuning, The visualization process helped me understand the distribution of features and how they influence predictions.