**Title: Predicting Instagram Likes Using Machine Learning**

**1. Introduction:**

In this mini-project, we aimed to predict the number of likes on Instagram posts using various machine learning models. The dataset contains 3785 posts with features such as the number of likes, followers, comments, and timestamps.

**2. Data Overview:**

The data was cleaned by removing rows with missing values and outliers. Outliers were found using Interquartile range. Along with that, added two new features – isWeekend, which determines the number of likes and also engagement percentage – which talks about the likes to follower count. Since I have used classification, I have divided the likes to 3 sections – Low , Medium and High.

**3. Model Training and Evaluation:**

Three models were trained:

- Random Forest

- Logistic Regression

- K-Nearest Neighbors

Evaluation was based on accuracy and error rate metrics.

**4. Results:**

- Random Forest: Accuracy = 96.52%, Error Rate = 3.48%

- Logistic Regression: Accuracy = 62.28%, Error Rate = 37.72%

- K-Nearest Neighbors: Accuracy = 75.44%, Error Rate =24.56%

**5. Conclusion:**

The Random Forest model performed the best, with the error rate of less than 4%.