**ML Assignment 5 – Random Forest, Naive Bayes, SVM**

**Crowdfunding Campaign Success Prediction**

The goal of this assignment is to develop and evaluate predictive classifier models to determine the success of crowdfunding campaigns based on a variety of campaign features. The dataset includes factors such as financial goals, actual funds raised, campaign duration, and backer engagement. By building and comparing three distinct machine learning models — Random Forest, Naive Bayes, and Support Vector Machine (SVM) — you will assess which model best predicts the success of crowdfunding efforts.

**Dataset Link:**

<https://raw.githubusercontent.com/ArchanaInsights/Datasets/refs/heads/main/crowdfunding_campaign.csv>

**Machine Learning Model Building Steps:-**

1. **Data Preprocessing:**
   1. **Load the Dataset:** Load the dataset and perform an initial exploration to understand its structure and content.
   2. **Handle Missing Values:** Check for and handle any missing values in the dataset if present.
   3. **Encode Categorical Features:** Convert categorical features into numerical format using appropriate encoding techniques.
   4. **Feature Selection:** Select the relevant features and the target variable for your analysis.
   5. **Data Splitting:** Split the dataset into training and testing sets (e.g., 75% training, 25% testing).
   6. **Feature Scaling**: Standardize or normalize numerical features using parameters derived from the training set, and apply the same scaling to the test set.
2. **Random Forest - Model Building and Evaluation:**
   1. **Model Building:** Build a Random Forest model using the training dataset. Experiment with different values for the number of trees (n\_estimators) in the forest and maximum depth (max\_depth).
   2. **Model Evaluation:** Evaluate the model on the testing dataset and report the accuracy score.
3. **Naive Bayes - Model Building and Evaluation:**
   1. **Model Building:** Build a Naive Bayes model using the training dataset. You can choose between Gaussian, Multinomial, or Bernoulli based on the nature of your features (e.g., GaussianNB() for continuous features).
   2. **Model Evaluation:** Evaluate the model on the testing dataset and report the accuracy score.
4. **Support Vector Machine (SVM) - Model Building and Evaluation:**
   1. **Model Building:** Build an SVM model using the training dataset. Experiment with different kernel functions (e.g., linear, RBF).
   2. **Model Evaluation:** Evaluate the model on the testing dataset and report the accuracy score.
5. **Comparison and Analysis:**
   1. Compare the accuracy scores of the three models.
   2. Discuss which model performed best and why, based on the accuracy scores and other relevant factors.
   3. Reflect on the strengths and weaknesses of each model in the context of this dataset.