**INXOL TECHNOLOGIES**

Team Lead

Muhammad Ikrma Sharaf

**Submitted by**

C1\_13\_Fatima Shareef

**Task 1:**

Gender Classification using Machine Learning Classifiers

Methodology

**Data Preparation:**

• The task starts by defining paths to the directories containing training and testing image data for both genders (male and female).

• Images are loaded using OpenCV, resized to a specified image size (64x64), and flattened into a 1D array.

**Creating Train and Test Data:**

Separate lists (train\_data\_female, train\_data\_male, test\_data\_female, test\_data\_male) are created to store images of different genders.

Images are loaded, resized, and added to their respective lists.

Train and test data are combined using concatenation techniques.

**Normalization and Labeling:**

Data is normalized by scaling it between 0 and 1.

Labels (target values) are created as 0 for female and 1 for male.

Train and test labels are concatenated and reshaped.

**Train-Test Split:**

The data is split into training and testing sets using train\_test\_split from scikit-learn.

**Classifier Evaluation:**

A list of classifier instances is defined, including Complement Naive Bayes, Decision Tree, Random Forest, Logistic Regression, SVM, K-Nearest Neighbors, Linear Discriminant Analysis, and Quadratic Discriminant Analysis.

For each classifier, the following steps are performed:

* The classifier is trained on the training data.
* Training time, training score, and test score are calculated .
* Confusion matrix is calculated and plotted as a heatmap.
* Classification report is printed, showing precision, recall, F1-score, and support for each class.

**Results Visualization:**

The code produces two bar graphs that visualize the training and testing accuracies of the classifiers. The first graph displays the accuracy achieved by each model, providing insights into their predictive performance. **Decision tree** and **random forest** gave the highest training accuracies while **SVC** gave the highest test accuracy .The second graph illustrates the training time required for each classifier, offering a comparison of computational efficiency among different models.

Among the presented classifiers, the Complement **Naive Bayes classifier** exhibits the shortest training time, indicating that it's computationally efficient and well-suited for rapid model training.

On the other hand, the **Random Forest** classifier takes relatively longer to train compared to others. This is expected, as Random Forest involves constructing multiple decision trees and combining their predictions, which can be more time-consuming.

**Conclusion:**

This code showcases a comprehensive approach to gender detection from facial images. By combining image processing techniques with machine learning algorithms, it demonstrates how to preprocess data, extract meaningful features, train models, and evaluate their performance. The visualizations and results provide valuable insights into the effectiveness and efficiency of different classifiers, aiding in the selection of the most suitable model for gender detection tasks.