# CSE/ECE 343 : Machine Learning Project Proposal
## Title : Face Recognition (Celebrities Identification)

Ayush Singal
ayush20365@iiitd.ac.in

Pranav Sharma
pranav20395@iiitd.ac.in

Tanish Gupta
tanis20344@iiitd.ac.in

## 1. Abstract

The project focuses on the technology of face recognition and its various applications in modern-day life. Face recognition technology is a rapidly advancing field that involves detecting, analyzing, and comparing facial features in digital images. The project provides an overview of how the technology works and its various applications, including security, surveillance, access control, and authentication.

The project explores the underlying algorithms and techniques used in face recognition systems, such as feature extraction, face detection, face alignment, and matching. It also covers various machine learning and deep learning techniques used in face recognition, such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Support Vector Machines (SVM), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN).

***keywords :*** Face recognition technology, Principal Component Analysis (PCA), Support Vector Machines (SVM), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN).

*The project discusses the ethical and legal issues associated with the use of face recognition technology, such as privacy concerns, bias, and misuse. It also highlights the ongoing research and development in the field, including new techniques for improving accuracy, robustness, and speed, as well as emerging applications in healthcare, entertainment, and education.*

Overall, the project aims to provide a comprehensive understanding of face recognition technology and its diverse applications, as well as its potential impact on society and the need for responsible and ethical use.

## 2. Timeline

A Tentative 6-7 week timeline :

Week 1-2: Data Collection
Week 2: Pre-processing and Data Visualization
Week 2-3: EDA (Class-imbalance, Correlations, Heatmaps, etc.)
Week 4: Multi-class Logistic Regression, MLP - Neural Network
Week 4-5: Convolutional Neural Networks
Week 5: Support Vector Machines (SVM)
Week 5-6: Decision Trees, Random Forest
Week 6: Analysis and performance of models
Week 6: Hyper-parameter Tuning, model over-fitting and under-fitting.
Week 7: Report writing

## 3. Individual Tasks

| Tasks | Team Member/s |
|---|---|
| Data Collection | Pranav and Tanish |
| Pre-processing and Data Visualization | Ayush and Tanish |
| Feature Extraction and Analysis | Ayush and Pranav |
| Analysis of Features (Selection , correlation, heatmaps. etc. . . .) | Ayush, Pranav and Tanish |
| Multi-class Logistic Regression, Support Vector Machines (SVM) | Pranav and Ayush |
| Decision Trees, Random Forest | Tanish and Ayush |
| LP - Neural Network, Convolutional Neural Networks | Pranav and Tanish |
| Hyperparameter Tuning , Check for Overfitting and underfitting of models, and selection of best model | All |
| Report Writing | All |

## 4. Introduction

Face recognition technology has become an essential aspect of security and surveillance systems in recent years. With the increasing need for reliable and accurate facial recognition, the development of a face recognition model using multiple machine learning models is a significant step in enhancing the technology. This project aims to develop a face recognition model using multiple machine learning models, including Support Vector Machines (SVM), Multi-Layer Perceptron (MLP), Convolutional Neural Networks (CNN), and Random Forest, with the CelebA dataset.

## 5. Background

Face recognition technology is a rapidly advancing field that involves detecting, analyzing, and comparing facial features in digital images or video frames. The development of a face recognition model using multiple machine learning models is a significant step towards improving the accuracy and reliability of face recognition technology.

## 6. Scope

The project will utilize the CelebA dataset, which consists of over 200,000 images of celebrities with various facial attributes. The project's scope will involve data pre-processing, feature extraction, model training, and testing, using multiple machine learning models.

## 7. Objectives

The primary objective of this project is to develop a robust and accurate face recognition model capable of recognizing faces in real-world scenarios accurately. Additionally, the project aims to analyze and compare the effectiveness of multiple machine learning models for developing a face recognition model.

## 8. Related Work

Face recognition is a popular application of machine learning, which involves training a model to recognize faces in images or videos :

1) We plan to explore research papers that propose methods to fuse SVM, MLP and CNN into a single real-time face recognition system for Real-time face recognition and enhancing the prediction accuracy [6].

2) Face recognition is categorized into three steps: face detection, face extraction and face identification. In this work, we create a face recognition application platform using Open-Computer- Vision (OpenCV), with a focus on the system's performance and accuracy [1].

3) We will also look into papers wherein face recognition application platform is created using openCV with a focus of the system's performance and accuracy [1].

## 9. Impact

The successful development of a reliable and accurate face recognition model using multiple machine learning models will have significant implications for security and surveillance systems, with potential applications in law enforcement, access control, and identification verification.

## 10. Materials and methods

The project will utilize the CelebA dataset for developing the face recognition model. The dataset will undergo pre-processing, including image resizing, normalization, and feature extraction. Multiple machine learning models, including SVM, MLP, CNN, and Random Forest, will be trained and tested using the pre-processed data. The evaluation metrics for the models will include accuracy, precision, recall, and F1 score.

## 11. Methodology and novelty

The methodology for this project involves data pre-processing, feature extraction, model training, and testing.

The novelty of the project lies in the use of multiple models such as SVM, MLP, CNN, and Random Forest for face recognition using the CelebA dataset. While these models have been used individually in face recognition research, combining them in this way allows for a more comprehensive and accurate approach to facial recognition. Additionally, the project aims to address ethical and legal concerns related to face recognition technology, which is an important and emerging issue in this field. Finally, the project discusses emerging applications of face recognition technology beyond traditional security and surveillance uses, such as in healthcare, entertainment, and education.

## 12. Evaluation metric(s)

The evaluation metrics for the developed face recognition models will include accuracy, precision, recall, and F1 score.

## 13. Our Goal

Overall, we aim to contribute to the development of face recognition technology by developing a reliable and accurate face recognition model using multiple machine learning models, with potential applications in security and surveillance systems with potential applications in law enforcement, access control, and identification verification.

## 14. References

1) Face Recognition System Using Computational Algorithms [1].

2) Facial Expression Classification Based on SVM, KNN and MLP Classifiers by Maiwan Bahjat Abdulrazzaq (2012) [2].

3) CNN: "DeepFace: Closing the Gap to Human-Level Performance in Face Verification" by Y. Taigman, M. Yang, M. Ranzato, and L. Wolf (2014) [3].

4) SVM: "Support vector machine-based face recognition: A comprehensive study" by Prof.Sheela Shankar, 2Dr.V.R Udupi[4].

5) Deep Face Recognition for Imperfect Human Face Images on Social Media using the CNN Method [5].

6) Real-time Face Recognition using SVM, MLP and CNN [6].