Insightful Identity Analysis: Detecting Age, Gender, and Ethnicity

Kunal Sharma
2021331
kunal21331@iiitd.ac.in

Sarvagya Kaushik
2021350
sarvagya21350@iiitd.ac.in

Vansh
2021363
vansh21363@iiitd.ac.in

1. Motivation

Age, gender, and ethnicity are crucial demographics in fields like marketing, healthcare, and social sciences. A machine learning model detecting these attributes can advance research and practical applications. It aids in healthcare for diagnosis and treatment planning. This project sharpens skills for real-world use, encompassing data prep, feature extraction, model training, and evaluation. It deepens understanding of machine learning algorithms and techniques.

2. Recent Works and Articles

Gender Prediction And Age Detection In Images Using OpenCV: [Medium Article]

This medium article has directions for the great computer vision library, OpenCV. This article is a tutorial where the author talks about the use of this machine learning framework.

Gender Detection Model using CNN — a complete guide: [Medium Article]

This article is about using Convolution Neural Networks in image classification. It is somewhat related to our project, but our one's more about finding all three categories (name, gender, and ethnicity). Before jumping into creating models, the author walks through all the procedures, such as converting images to numbers and data cleaning.

Gender and age prediction from real time facial images using CNN: [Article]

The above research article gives deeper insight into CNN and its use cases. It talked about the use case of the project and various methods to enhance the model's accuracy. Overall, it provides deeper insight into the problem and ways to tackle it.

3. Timeline for Project

Data Analysis:
(August 30 - September 30)
We'll be carefully analyzing data using various charts and graphs to gain insights about our data. We'll try to find relations among features and also identify data points that can corrupt our model.

Data Cleaning:
(September 30 - October 15)
After analyzing data and identifying outliers, we'll clean our data to prepare it for building useful models. We'll also convert given image data into matrices with pixel intensities as entries.

Model Preparation:
(October 15 - November 15)
During this month, we'll build models using various frameworks and libraries. We'll experiment with different approaches and create multiple models.

Optimization and Finalization:
(November 15 - End sems)
This period will involve code cleanup, optimization, and final adjustments. We'll also use this time if we don't meet previous deadlines.

4. Individual Tasks

Kunal Sharma:
a. Data Analysis.
b. Finding important insights/patterns in data.
c. Hyperparameter tuning.
d. Trying out different models and documentation.

Sarvagya Kaushik:
a. Model Preparation.
b. Data cleaning.
c. Optimization and Finalization.
d. Documentation work.

Vansh:
a. Data Analysis.
b. Documentation work.
c. Optimization and Finalization.
d. Experimenting with different ML frameworks.

5. Final Outcome

Through this project, we will have gained a solid understanding of machine learning techniques, such as artificial neural networks(ANN) and convolutional neural networks (CNN). We will have gained practical experience with machine learning frameworks, such as TensorFlow, PyTorch, or Keras. Through this project, we will have gained experience in training machine learning models using large datasets. We will have learned how to evaluate the performance of your models using appropriate evaluation metrics, such as accuracy, precision, recall, or F1 score. By working on this project, we will have gained insights into the real-world applications of age, gender, and ethnicity detection. These applications can range from facial recognition systems to personalised marketing and demographic analysis.