



INTRODUCTION TO DATA SCIENCE

ASSIGNMENT # 03

MACHINE LEARNING

AROoba MASOOD

FA20-BSE-092

COMSATS Institute of Information Technology

Department of Computer Science

23 November 2023

Provide responses to the following questions about the dataset.

1. How many instances does the dataset contain?

The dataset contains total of 110 datasets.

2. How many input attributes does the dataset contain?

The dataset contains 7 input attributes:

- Height
- Weight
- Beard
- hair_length
- shoe_size
- scarf
- eye_color.

3. How many possible values does the output attribute have?

The output attribute 'gender' has 2 possible values: male and female.

4. How many input attributes are categorical?

The categorical input attributes are: beard, hair_length, scarf, eye_color. So there are 4 categorical attributes.

5. What is the class ratio (male vs female) in the dataset?

There are 62 male instances and 48 female instances. So the class ratio of male to female is 62:48

