FA20-BSE-057(A) ESHA NAEEM

Assignment 3(IDS)

Question 1:

How many instances does the dataset contain?

There are 110 instances in the dataset.

How many input attributes does the dataset contain?

There are 7 input attributes: height, weight, beard, hair_length, shoe_size, scarf, and eye_color.

How many possible values does the output attribute have?

The output attribute is "gender," and it seems to have two possible values: "male" and "female."

How many input attributes are categorical?

Categorical attributes are those that can take on a limited, and usually fixed, number of possible values. In your dataset, the categorical attributes are "beard," "hair_length," "scarf," and "eye_color."

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

the number of instances for each gender.

Number of males: 62

Number of females: 48

Class ratio (male vs female): 62:48

Question 2:

PART3

Beard: Explanation: The presence or absence of a beard is often strongly associated with gender, as beards are more commonly associated with males. This attribute can serve as a significant indicator for the model in predicting gender.

Height: Explanation: While there is overlap in the height distribution between genders, on average, males tend to be taller than females. Therefore, height can be a useful feature for predicting gender. However, it's important to note that relying solely on height for gender prediction might lead to misclassifications due to the natural variation in height within each gender.

Note: rest parts explanation in shows in code

Question 4:

NEW Instances

112	71	151	no	medium	40	yes	black	female
113	72	152	no	medium	42	yes	brown	female
114	49	153	no	long	37	no	black	female
115	50	155	no	long	49	yes	brown	female
116	67	165	no	medium	45	yes	blue	female
117	43	166	yes	short	36	no	black	male
118	70	167	yes	short	37	no	brown	male
119	72	168	yes	short	38	no	brown	male
120	68	120	yes	short	49	no	black	male
121	69	140	yes	short	48	no	black	male