Assignment 5 - Question 3

Machine Learning / AI 20 points

DESCRIPTION

Q3: After performing the analysis from the previous question, derive a new column called "AdultOrChild" having categorical values as "Adult" or "Child" derived from Age column

Hint: A person having Age >=18 is an "Adult" and the one having Age < 18 is a "Child".

1. Find its relation with the "Survived" Column and print the total number of survivors.

Example:

If Total survived children: 100, Total survived adults: 200

Output: 300

2. Choose features to create a Classification model and predict the survived category

For the above prediction create a Confusion matrix for the model built by you and print the sum of all the elements of a matrix

******NOTE: You should create the confusion matrix for the test data not the training data.***

Example: If the Confusion Matrix is [2 2

2 2]

(2+2+2+2)

Output: 8

Hint: Use Logistic Regression as the classification model

3. Use confusion matrix to print the accuracy of the model

Example: (2+2)/8*100

Output: 50

******NOTE: You should check the accuracy for the test data not the training data.***

Final Output Sample:

	A
1	100
2	200
3	300

image.png 3.67 KB

NOTE: Here, 100, 200 and 300 are the answer of 1st, 2nd and 3rd question respectively.

Output Format:

- Perform the above operations and write (written above as **print**) your output to a file named **output.csv**, which should be present at the location **/code/output/output.csv**
- **output.csv** should contain the answer to **each question** on consecutive rows.

******NOTE: For all the questions the numerical values saved in output.csv file should be in integer format with no decimals.***

DATASETS

- Training dataset