Program that implements logistic regression with gradient descent algorithm.

We have data extracted from images that were taken from genuine and forged banknote-like specimens as training data for logistic regression.

|  |  |  |  |
| --- | --- | --- | --- |
| **Data Set Characteristics:** | Multivariate | **Number of Instances:** | 1372 |
| **Attribute Characteristics:** | Real | **Number of Attributes:** | 5 |
| **Associated Tasks:** | Classification | **Missing Values?** | No |

Attribute Information:

1. variance of Wavelet Transformed image (continuous)
2. skewness of Wavelet Transformed image (continuous)
3. curtosis of Wavelet Transformed image (continuous)
4. entropy of image (continuous)
5. class (integer, 0: forged 1: genuine)

What the code does

1. Implements logistic regression with gradient descent
2. Uses Cross-Validation to calculate the error rate. In Cross-Validation, randomly split the training data into ratio 90:10. Use the 90% as training data and predict labels on the remaining 10%. Repeat 10 times and average the error rate to get your Cross-Validation estimate.

Source of the Dataset: <http://archive.ics.uci.edu/ml/datasets/banknote+authentication>

Find more datasets on this website: <http://archive.ics.uci.edu/ml/datasets.html>