|  |  |
| --- | --- |
| **banknote authentication Data Set** *Download*: [Data Folder](https://archive.ics.uci.edu/ml/machine-learning-databases/00267/), [Data Set Description](https://archive.ics.uci.edu/ml/datasets/banknote+authentication)  **Abstract**: Data were extracted from images that were taken for the evaluation of an authentication procedure for banknotes. |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Data Set Characteristics:** | Multivariate | **Number of Instances:** | 1372 | **Area:** | Computer |
| **Attribute Characteristics:** | Real | **Number of Attributes:** | 5 | **Date Donated** | 2013-04-16 |
| **Associated Tasks:** | Classification | **Missing Values?** | N/A | **Number of Web Hits:** | 340046 |

**Source:**

Owner of database: Volker Lohweg (University of Applied Sciences, Ostwestfalen-Lippe, volker.lohweg **'@'** hs-owl.de)  
Donor of database: Helene DÃ¶rksen (University of Applied Sciences, Ostwestfalen-Lippe, helene.doerksen **'@'** hs-owl.de)  
Date received: August, 2012

**Data Set Information:**

Data were extracted from images that were taken from genuine and forged banknote-like specimens. For digitization, an industrial camera usually used for print inspection was used. The final images have 400x 400 pixels. Due to the object lens and distance to the investigated object gray-scale pictures with a resolution of about 660 dpi were gained. Wavelet Transform tool were used to extract features from images.

**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)

We will fit a linear function to this dataset that predicts the value of the fifth attribute using the first 4 attributes; we will determine the R\_squared, the coefficient of determination for the obtained regression function, which takes values in [0,1]. If the fit is good; e.g. above 0.7, we can use the attribute coefficients to identify important attributes; if the fit is poor, as the model is bad one has to be cautious to draw conclusions about the importance of attributes based on the regression coefficients.

[Coefficient of determination - Wikipedia](https://en.wikipedia.org/wiki/Coefficient_of_determination)

[8 Tips for Interpreting R-Squared - Displayr](https://www.displayr.com/8-tips-for-interpreting-r-squared/#:~:text=8%20Tips%20for%20Interpreting%20R-Squared%201%20Don%27t%20conclude,R-Squared%20to%20compare%20models.%20...%20More%20items...%20)