**Dataset Description**

**# Load datasets**

**def load\_datasets():**

**# Load training data**

**train\_df = pd.read\_csv('bankruptcy\_Train.csv')**

**test\_df = pd.read\_csv('bankruptcy\_Test\_X.csv')**

**File descriptions**

* **bankruptcy\_Train.csv** - the training set with 64 predictors and 1 target variable
* **bankruptcy\_Test\_X.csv** - the test set with ID and 64 predictors
* **bankruptcy\_sample\_submission.csv** - the sample submission with ID and the predicted probability of firm bankruptcy

**Data fields**

* **attr1** - net profit / total assets
* **attr2** - total liabilities / total assets
* **attr3** - working capital / total assets
* **attr4** - current assets / short-term liabilities
* **attr5** - [(cash + short-term securities + receivables - short-term liabilities) / (operating expenses - depreciation)] \* 365
* **attr6** - retained earnings / total assets
* **attr7** - EBIT / total assets
* **attr8** - book value of equity / total liabilities
* **attr9** - sales / total assets
* **attr10** - equity / total assets
* **attr11** - (gross profit + extraordinary items + financial expenses) / total assets
* **attr12** - gross profit / short-term liabilities
* **attr13** - (gross profit + depreciation) / sales
* **attr14** - (gross profit + interest) / total assets
* **attr15** - (total liabilities \* 365) / (gross profit + depreciation)
* **attr16** - (gross profit + depreciation) / total liabilities
* **attr17** - total assets / total liabilities
* **attr18** - gross profit / total assets
* **attr19** - gross profit / sales
* **attr20** - (inventory \* 365) / sales
* **attr21** - sales (n) / sales (n-1)
* **attr22** - profit on operating activities / total assets
* **attr23** - net profit / sales
* **attr24** - gross profit (in 3 years) / total assets
* **attr25** - (equity - share capital) / total assets
* **attr26** - (net profit + depreciation) / total liabilities
* **attr27** - profit on operating activities / financial expenses
* **attr28** - working capital / fixed assets
* **attr29** - logarithm of total assets
* **attr30** - (total liabilities - cash) / sales
* **attr31** - (gross profit + interest) / sales
* **attr32** - (current liabilities \* 365) / cost of products sold
* **attr33** - operating expenses / short-term liabilities
* **attr34** - operating expenses / total liabilities
* **attr35** - profit on sales / total assets
* **attr36** - total sales / total assets
* **attr37** - (current assets - inventories) / long-term liabilities
* **attr38** - constant capital / total assets
* **attr39** - profit on sales / sales
* **attr40** - (current assets - inventory - receivables) / short-term liabilities
* **attr41** - total liabilities / ((profit on operating activities + depreciation) \* (12/365))
* **attr42** - profit on operating activities / sales
* **attr43** - rotation receivables + inventory turnover in days
* **attr44** - (receivables \* 365) / sales
* **attr45** - net profit / inventory
* **attr46** - (current assets - inventory) / short-term liabilities
* **attr47** - (inventory \* 365) / cost of products sold
* **attr48** - EBITDA (profit on operating activities - depreciation) / total assets
* **attr49** - EBITDA (profit on operating activities - depreciation) / sales
* **attr50** - current assets / total liabilities
* **attr51** - short-term liabilities / total assets
* **attr52** - (short-term liabilities \* 365) / cost of products sold)
* **attr53** - equity / fixed assets
* **attr54** - constant capital / fixed assets
* **attr55** - working capital
* **attr56** - (sales - cost of products sold) / sales
* **attr57** - (current assets - inventory - short-term liabilities) / (sales - gross profit - depreciation)
* **attr58** - total costs /total sales
* **attr59** - long-term liabilities / equity
* **attr60** - sales / inventory
* **attr61** - sales / receivables
* **attr62** - (short-term liabilities \*365) / sales
* **attr63** - sales / short-term liabilities
* **attr64** - sales / fixed assets
* **class** - the response variable Y: 0 = did not bankrupt; 1 = bankrupt