MINI - HACKATHON COMPETITION 2018

FINAL ROUND

I. Problem

In a financial industry, a credit rating is an evaluation of the credit risk of a prospective debtor, predicting their ability to pay back the debt, and an implicit forecast of the likelihood of the debtor defaulting. In this competition, we are facing with some companies in Construction Engineering industry which are needed to measure their default. We provide you several company data and their status over years, the data is extracted from their financial statements.

II. Dataset

The dataset contains 2 files:

- `train.csv`: list all data points for training,

- `test.csv`: list all data points for testing,

II.1. Features

Both train set and test one have the following schema (see Table 1)

Field	Type	Description
Gross Profit	float	the profit a company makes after deducting the costs associated with making and selling its products, or the costs associated with providing its services
Revenue	float	the amount of money that a company actually receives during a specific period, including discounts and deductions for returned merchandise
Pretax Income	float	a company's earnings after all operating expenses, including interest and depreciation, have been deducted from total sales or revenues, but before income taxes have been subtracted
Total Assets	float	the sum of all current and noncurrent assets and must equal the sum of total liabilities and stockholders' equity combined
Total Equity	float	total equity is equal to a firm's total assets minus its total liabilities
Total Liabilities	float	total liabilities refer to the aggregate of all debts an individual or company is liable

Total Current Assets	float	the value of all assets that can reasonably expect to be converted into cash within one year. Current assets include cash and cash equivalents, accounts receivable, inventory, marketable securities, prepaid expenses and other liquid assets that can be readily converted to cash
Total Current Liabilities	float	a company's debts or obligations that are due within one year
Inventories	float	the raw materials, work-in-process products and finished goods that are considered to be the portion of a business's assets that are ready or will be ready for sale
Interest Expense	float	the cost incurred by an entity for borrowed funds
EBITDA	float	EBITDA stands for earnings before interest, taxes, depreciation and amortization
Short-Term Borrowings	float	all indebtedness in respect of borrowed money maturing on demand or within one year from the date of the creation thereof
Cash & Near Cash Items	float	Cash and cash equivalents are either cash or assets that are readily convertible into cash
Cost of Revenue	float	the total cost of manufacturing and delivering a product or service
Accounts & Notes Receivable	float	receivables represent funds that a company is counting on collecting but that are as yet still outstanding
symbol	int	an encoded symbol of company
year	int	a year of this financial data point
default	int	whether a company in default at the end of the year respectively, 1 is default and 0 is non-default

Table 1: List of features

II.2. Train set

The train set has 377 data points with 18 attributes which are described in the section above.

II.3.   Test set

The test set has 168 data points and doesn't contains column *default*.

III.   Performance Metric

There is no default metric, it means that you can propose and use any metric(s) suitable with your approaches.

IV.   Requirements

Our main goal is to predict a company will be default or not, but you are free to propose and do other analysis on this dataset.

At the end of this competition, please submit your prediction on the test set via one csv file with a format below.

```
index , default
0 ,1
1 ,1
2 ,0
3 ,0
4 ,0
. . .
```

V.   Report

- The report should be presented in good format and English usage,

- The report should have clear structure to deliver your work efficiently.