PROJECT : PYTHON FOR DATA ANALYSIS

By Suvin Sasikumar
and Chanelle Wea

DIA 7

GOALS

The objective of this presentation are :

the ins and outs of the problem

thoughts of the asked questions

the different variables created

how the problem fits in the context of the study

How THE PRoBLEM FiTS iN THE CONTEXT OF STuDY

When you have a large dataset that's full of interesting insigths but you're not sure where to starts exploring it. Pandas and Python can be a good help. We are able to slice a large dataset down into manageable parts and glean insight front that information. We can visualize our data with plots, discover and handle incorrect data and missing values.

This semester, we saw the basics of Python, Numpy, Pandas, Matplotlib, Seaborn, Web Scraping, Supervided Learning and Tensorflow. All these libraries helped us to analyze more precisely a dataset.

DATA SET INFORMATION :

The dataset is about bankruptcy prediction of Polish companies. The data was collected from Emerging Markets Information Service, which is a database containing information on emerging markets around the world.

We decided to focus on the data that contains financial rates from 1st year of the forecasting period and corresponding class label that indicates bankruptcy status after 5 years. The data contains 7027 instances (financial statements), 271 represents bankrupted companies, 6756 firms that did not bankrupt in the forecasting period.

INS AND OUTS

Observations of the data, data analysis and preprocessing, data-visualization, correlation matrix

Data Modeling : Prediction Model for column X65

The X65 column contains 0 if the company has not bankrupt, 1 if it has.

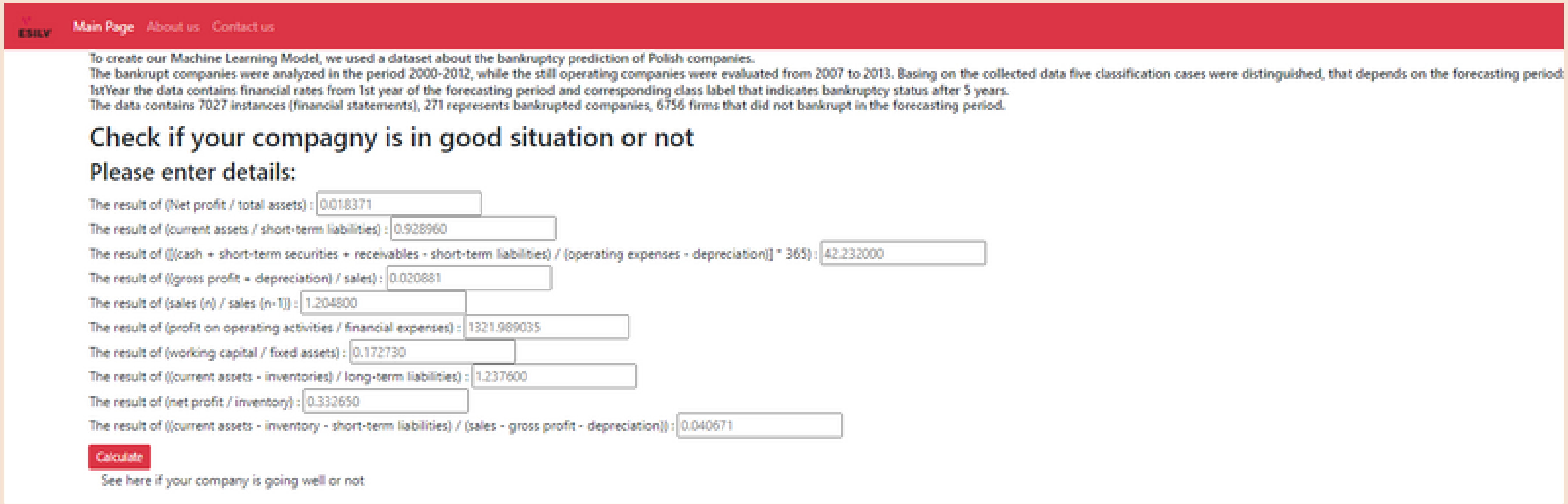
Random Forest : 100% | Logistic Regression : 96%
Decision Tree : 94.7% | Bagging : 96.3%
Gaussian Naive Bayes : 7.6% | Gradient Boosting : 94%
KNN : 96%

THE DiFFERENTS VARiABLES CREATED

- Parts of Bankruptcy company
- Preprocessing
- Data Analysis
- Correlation Matrix
- Data Visualisation
- Bankruptcy features
- Prediction Model
- Comparison of the Models

OUR APi

We created a
Django Api.



To create our Machine Learning Model, we used a dataset about the bankruptcy prediction of Polish companies.
The bankrupt companies were analyzed in the period 2000-2012, while the still operating companies were evaluated from 2007 to 2013. Basing on the collected data five classification cases were distinguished, that depends on the forecasting period:
1stYear the data contains financial rates from 1st year of the forecasting period and corresponding class label that indicates bankruptcy status after 5 years.
The data contains 7027 instances (financial statements), 271 represents bankrupted companies, 6756 firms that did not bankrupt in the forecasting period.

Check if your compagny is in good situation or not

Please enter details:

The result of (Net profit / total assets) : 0.018371

The result of (current assets / short-term liabilities) : 0.928960

The result of ([(cash + short-term securities + receivables - short-term liabilities) / (operating expenses - depreciation)] * 365) : 42.232000

The result of ((gross profit + depreciation) / sales) : 0.020881

The result of (sales (n) / sales (n-1)) : 1.204800

The result of (profit on operating activities / financial expenses) : 1321.989035

The result of (working capital / fixed assets) : 0.172730

The result of ((current assets - inventories) / long-term liabilities) : 1.237600

The result of (net profit / inventory) : 0.332650

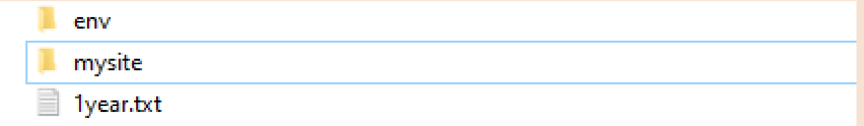
The result of ((current assets - inventory - short-term liabilities) / (sales - gross profit - depreciation)) : 0.040671

Calculate

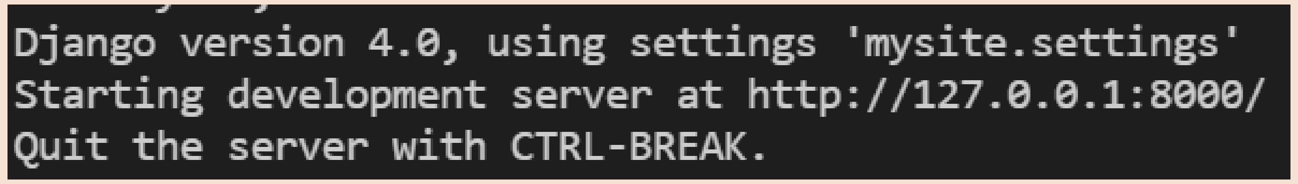
See here if your company is going well or not

How to use the API ?

- Open Visual Studio Code, and open the file "my site"



- open a new terminal command line
- write this code on the terminal C:\Users\Desktop\Django ML\mysite> python manage.py runserver
- you will obtain a link to a html website : http://127.0.0.1:8000/
- copie and paste it to an internet navigator.



```
Django version 4.0, using settings 'mysite.settings'
Starting development server at http://127.0.0.1:8000/
Quit the server with CTRL-BREAK.
```

- Now you can use our API ! You have to enter all the details asked, and our machine Learning model will predict if your compagny is going to bankrupt or not !