Applying Machine Learning To Financial Risk Management Using IBM Watson

1. INTRODUCTION

1.1 Overview

Financial institutions need to continually weigh the risks of their transactions, and they determine their risk level through credit scoring. Leading up to the 2008-09 financial crisis, almost all large banks used credit scoring models based on statistical theories; that crisis, largely brought about by underestimating risk, proved the need for better accuracy in their scoring. The combination of increased requirements and the development of advanced new technologies has given rise to a new era: credit scoring using machine learning.

1.2 Purpose

With this project, we can focus on detecting, managing and securing exposure to various risks arising from the use of financial services. First, the approach is more efficient and allows the bank to do more with less manpower. Second, it is more effective. The fact that incidents can be detected earlier allows the bank to prevent them from spiraling out of control. Third, the system is adaptive. Humans have a great capacity to adapt to controls imposed on them.

2. LITERATURE SURVEY

2.1 Existing problem

Three commonly used approaches to measure financial risks are regression analysis, Value-at-Risk analysis, and scenario analysis.

a. Regression analysis

Ease of Use- Simple

Uses- Reducing exposure to specific risk factors e.g., exchange rate movements, Determining hedging strategies.

Advantages- Excel-based, Easy to understand

Disadvantages- Regression equation may not be stable over time making the results movements unreliable.

b. Value-at-Risk analysis

Ease of Use- Potentially complex, requiring good statistical understanding

Uses- Enhances understanding of a wide range of risks covering liquidity, cash flows, portfolio values, credit, etc. Can be used as a risk control tool.

Advantages- Easy to understand, Gives a sense of the likelihood of a given scale of losses.

Disadvantages- No idea of the potential scale of losses in excess of VaR. May give a false sense of security because it does not capture extreme scenarios.

c. Scenario analysis

Ease of Use- Simple

Uses- 'What if' analyses, Crisis planning

Advantages- Highly flexible, Easy to understand

Disadvantages- Likelihood of alternative scenarios may not be easily assessed. Specification of scenarios is subjective.

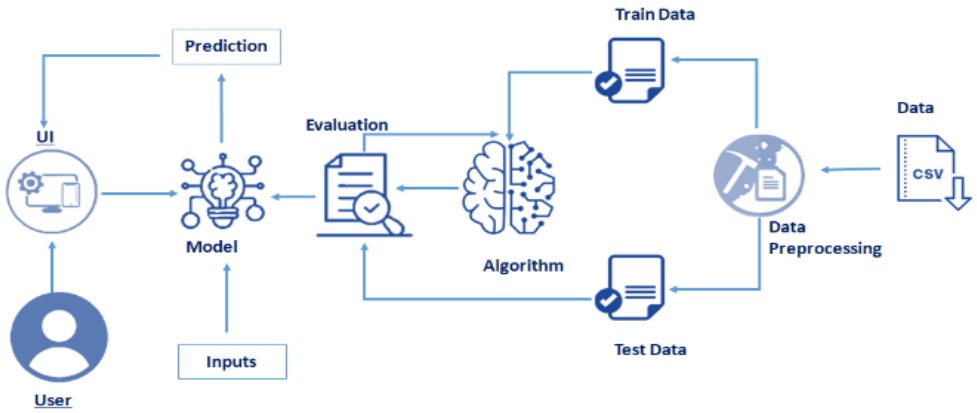
2.2 Proposed solution

The input values are inputted first, and the model then analyses the entered input values. Following input analysis by the model, the prediction is presented. We must perform each of the four steps outlined below in order to achieve this

- Data Collection- where we collect the dataset or create the dataset
- Data Preprocessing- where we Import the Libraries and dataset. Check for Null Values and perform data Visualization, Label Encoding and OneHot Encoding. And then we Split the data into Train and Test, and perform feature Scaling.
- Model Building- where we Import the model building libraries. First we initialize the model, and then Train and test it. After which the evaluation of model is done and saved.
- Application Building- where we will be building a web application that is integrated to the model we built. A UI is provided for the uses where he has to enter the values for predictions. The entered values are given to the saved model and prediction is showcased on the UI. We will be building HTML Pages and server side script here.

3. THEORITICAL ANALYSIS

3.1 Block diagram



3.2 Hardware / Software designing

Software requirements-

- Anaconda navigator
- Python packages:

  pip install numpy

  pip install pandas

  pip install matplotlib

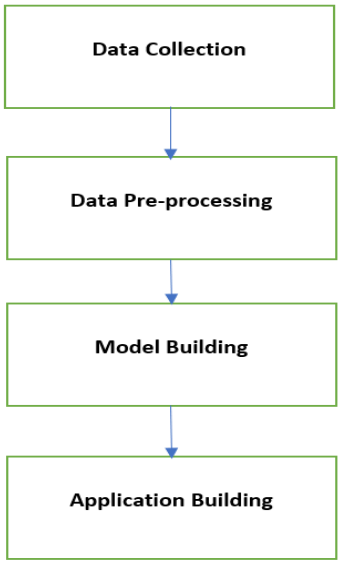
  pip install scikit-learn

  pip install Flask
- Jupyter notebook
- Python flask

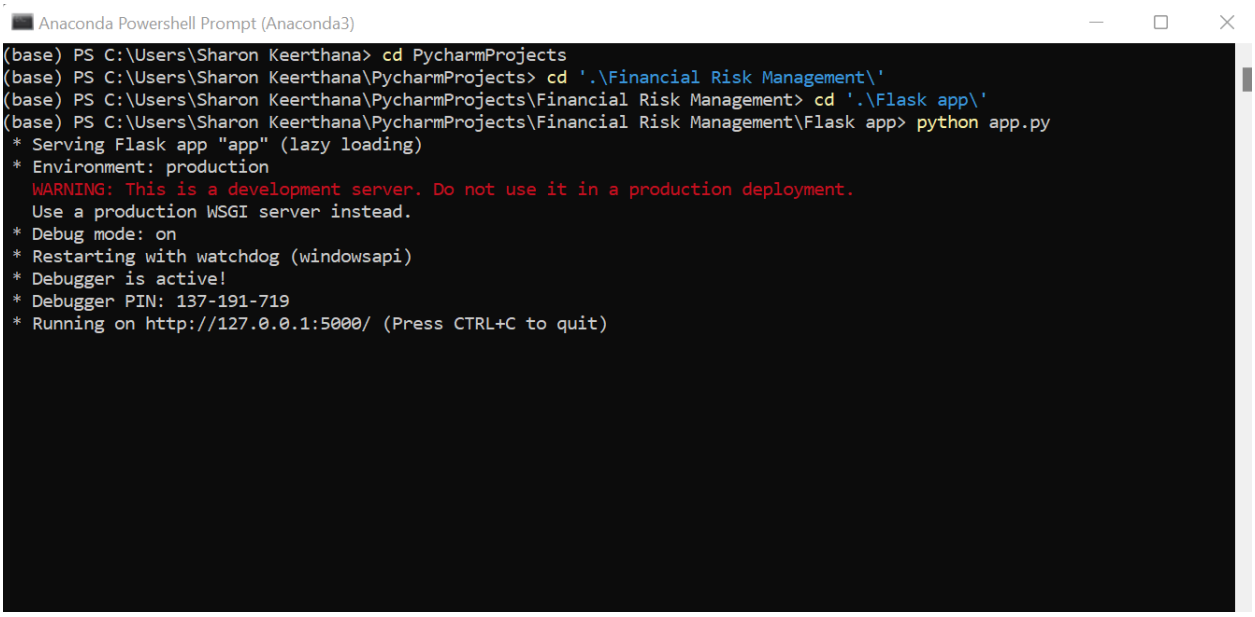
4. EXPERIMENTAL INVESTIGATIONS

The dataset contains of 1000 entries, each entry represents a person who takes a credit by a bank. User interacts with the UI (User Interface) to enter the input values such as age, gender, employment type, checking account, saving account, type of housing, purpose, credit amount and duration. Entered input values are analyzed by the model which is integrated. Once model analyses the input the prediction is showcased on the UI whether each person is classified as good or bad credit risks.

5. FLOWCHART



6. RESULT



127.0.0.1:5000

FINANCIAL RISK ASSESMENT

Financial institutions need to continually weigh the risks of their transactions, and they determine their risk level through credit scoring. Leading up to the 2008-09 financial crisis, almost all large banks used credit scoring models based on statistical theories; that crisis, largely brought about by underestimating risk, proved the need for better accuracy in their scoring. The combination of increased requirements and the development of advanced new technologies has given rise to a new era: credit scoring using machine learning.

Click me!

127.0.0.1:5000/assesment?

FINANCIAL RISK ASSESMENT

Enter the Age

Select the Gender

Choose employment type

Choose Checking Account

Select the type of Housing

Choose Savings Account ...

Select the Purpose

Enter the Credit Amount

Enter the Duration

Submit

127.0.0.1:5000/assesment?

FINANCIAL RISK ASSESMENT

41

Male

Unskilled / Resident

Rich

Rent

Moderate

Business

345009

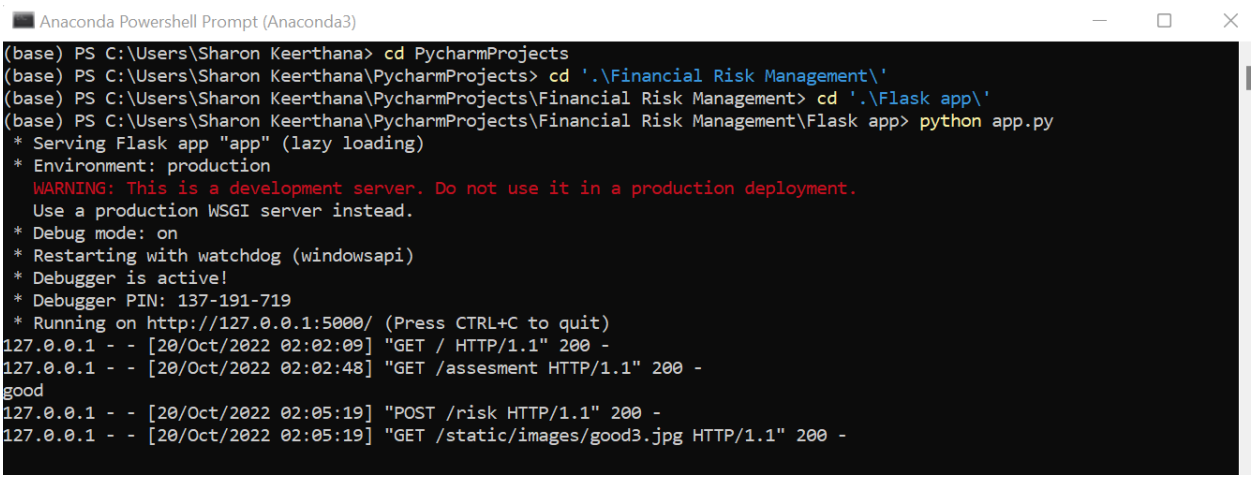
7

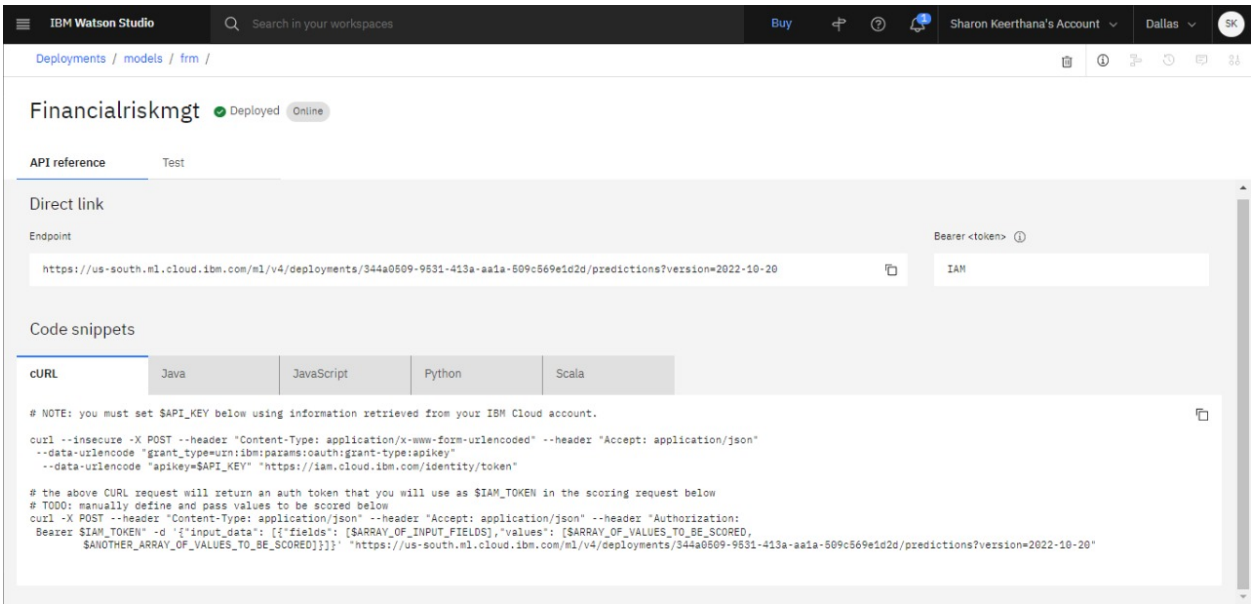
Submit

127.0.0.1:5000/risk

FINANCIAL RISK ASSESMENT

THE FINANCIAL RISK ASSESED IS GOOD





7. ADVANTAGES & DISADVANTAGES

Advantages-

- Forecasts Probable Issues
- Avoiding Catastrophic Events
- Enables Growth
- Helps to Stay Competitive
- Business Process Improvement
- Enables Better Budgeting

Disadvantages-

- Not Suitable For All Organizations

- Expensive
- Training Costs
- Loss of Focus Due to Automation
- Data Security Issue

8. APPLICATIONS

- Operational Risk-

Indirect or direct loss caused by failed or inadequate internal people, system, processes or external events. It includes risk types such as security, legal, fraud, environmental and physical risks (major power failures, etc.). Operational risks are not revenue driven, incurred knowingly or capable of being completely eliminated. Operational risks can be managed to acceptable levels of risk tolerance by determining the costs of a proposed improvement against its benefits.

- Foreign Exchange Risk-

Incurred when a financial transaction is made in a currency other than the operating currency of a business, arising as a result of unfavorable changes in the exchange rate between the two. An aspect of Foreign Exchange Risk is Economic Risk or Forecast Risk; the degree to which an organisation's product or market value is affected by unexpected exchange-rate fluctuations.

- Credit Risk-

Incurred if a borrower defaults on their debts or outstanding payments. With borrowed money, in addition to the loss of principal, additional factors such as loss of interest, increasing collection costs etc. must be taken into account when establishing the extent of the Credit Risk. Financial analysts use Yield Spreads as means to determine Credit Risk levels in a market. Ways of mitigating Credit Risk is to run a credit check, purchase insurance, hold assets as collateral or guarantee the debt by a third-party.

- Reputational Risk-

The loss of social capital, market share or financial capital arising from damage to an organisation's reputation. Reputation Risk is very difficult to predict or realize financially, as Reputation is an intangible asset. It is however intrinsically tied to Corporate Trust and is the reason why Reputation damage can hurt an organization financially through consumer boycotts. In extreme cases, Reputational Risk can even lead to corporate bankruptcy. For this reason, more organizations are dedicating assets and resources to better manage their reputation.

9. CONCLUSION

- We were able to understand the problem to classify if it is a regression or a classification kind of problem.
- We were able to know how to pre-process / clean the data using different data preprocessing techniques.

- We were able to analyze or get insights of data through visualization.
- We were able to apply different algorithms according to dataset and based on visualization.
- We were able to know how to find the accuracy of the model.
- We were able to know how to build a web application using the Flask framework.

We used many machine learning algorithms and out of those, logistic regression gave the maximum accuracy.

10. FUTURE SCOPE

The proposed model can be improved further to make it even more efficient and accurate. These can be done through the following ways:

- For complex tasks that suffer from statistically problematic data attributes such as non-stationary or high amount of noise, robust learning methods may be implemented. Since these problematic data attributes (high noise or low signal-to-noise ratio) are not only specific to the finance domain, there is an active stream of research across different domains to learn from extremely noisy data.
- High-dimensional parametric models such as Neural networks are known to accurately identify latent correlations among features and labels. Most of the current work in literature focuses on applying these techniques mainly for prediction purposes, but very few attempt to find the causal relationship between features and model prediction. However, for many sensitive tasks such as credit approval decisions or insurance pricing, causal explanation of model predictions can be of significant importance to both model developers and regulators
- A standard technique for achieving Differential privacy is the Laplace mechanism, which adds noise without deteriorating final model performance. Whereas federated learning techniques mainly focus on preventing sharing private data across multiple parties, differential privacy can add one more layer to overall model security by ensuring that the federated model weights cannot be used to identify private information.

11. BIBILOGRAPHY

[1] W.-Y. Lin, Y.-H. Hu, and C.-F. Tsai, ‘‘Machine learning in financial crisis prediction: A survey,’’ IEEE Trans. Syst., Man, Cybern. C, Appl. Rev., vol. 42, no. 4, pp. 421–436, Jul. 2012.
[2] P. O. J. Kelliher, D. Wilmot, J. Vij, and P. J. M. Klumpes, ‘‘A common risk classification system for the actuarial profession,’’ Brit. Actuarial J., vol. 18, no. 1, p. 91–121, 2013.

[3] S. Emerson, R. Kennedy, L. O’Shea, and J. O’Brien, ‘‘Trends and applications of machine learning in quantitative finance,’’ in Proc. 8th Int. Conf. Econ. Finance Res. (ICEFR), 2019.

[4] K. A. Horcher, Essentials of Financial Risk Management, vol. 32. Hoboken, NJ, USA: Wiley, 2011.

[5] P. Christoffersen, Elements of Financial Risk Management. New York, NY, USA: Academic, 2011

[6] D. Bisias, M. Flood, A. W. Lo, and S. Valavanis, ‘‘A survey of systemic risk analytics,’’ Annu. Rev. Financial Econ., vol. 4, no. 1, pp. 255–296, Oct. 2012.

[7] J. De Spiegeleer, D. B. Madan, S. Reyners, and W. Schoutens, ‘‘Machine learning for quantitative finance: Fast derivative pricing, hedging and fitting,’’ Quant. Finance, vol. 18, no. 10, pp. 1635–1643, Oct. 2018.

[8] T.-M. Choi, H. K. Chan, and X. Yue, ‘‘Recent development in big data analytics for business operations and risk management,’’ IEEE Trans. Cybern., vol. 47, no. 1, pp. 81–92, Jan. 2017.

[9] A. Bahrammirzaee, ‘‘A comparative survey of artificial intelligence applications in finance: Artificial neural networks, expert system and hybrid intelligent systems,’’ Neural Comput. Appl., vol. 19, no. 8, pp. 1165–1195, Nov. 2010.

[10] G. Kou, X. Chao, Y. Peng, F. E. Alsaadi, and E. Herrera-Viedma, ‘‘Machine learning methods for systemic risk analysis in financial sectors,’’ Technol. Econ. Develop. Economy, vol. 25, no. 5, pp. 716–742, May 2019.

APPENDIX

A. Source Code

SI-GuidedProject-326050-1665127367/internship.ipynb at main · smartinternz02/SI-GuidedProject-326050-1665127367 (github.com)