Investigating Super learner for credit Risk modeling in Mortgage Scenario

Lalit Aggarwal
Research Scholar
Reva University
Bangalore, India
lalitaggarwal.ba05@reva.edu.in

Sanjeev Jha
Research Scholar
Reva University
Bangalore, India
sanjeev.ba05@reva.edu.in

Dr. J.B. Simha
Chief Mentor
Reva University
Bangalore, India
jb.simha@reva.edu.in

In the present industry Credit risk analysis is very important for the organization's business as well as its reputation in the market. In general, credit risk modelling is a method that lenders employ to assess the degree of credit risk involved in making a loan to a borrower. When a business or individual borrower doesn't fulfil their loan obligations, credit risk develops. It is the likelihood that a lender won't get the principal and interest payments needed to pay off a loan that has been issued to a borrower. Credit risk will affect the lender's cash flow and raise collection expenses since it may be required to work with a debt collection agency to enforce the collection. The lender may suffer a loss of all or a portion of the loan that was given to the borrower, depending on the circumstances.

The turnaround time for delivering a statistical model in the general machine learning approach is somewhat long. Creating machine learning models that can reliably anticipate the outcomes is always difficult and time-consuming, even for expert data scientists. There are several stages to the intricate workflow that goes into machine learning models. While Automatic Machine Learning (AutoML) is more concerned with gathering data and making predictions. The AutoML platform will abstract every step that occurs between these two phases. AutoML models require extremely little time to develop models and tune hyperparameters. Prediction can also be extremely well conveyed with the use of various outcome interpreter approaches.

The objective of this project is to investigate Super learners for Credit Risk modeling in Mortgage scenarios, with the help of Automation Machine learning. Different super learners can be defined as the ensemble models of different base models and investigated against the credit risk dataset. For explaining the prediction of the super learner different result interpretation techniques i.e. SHAP, PDP and ICE have been used.

With the Use of H2O AutoML and the credit risk data, multiple machine learning models have been used, which comprise 25 models including Ensemble machine learning models “StackEnsemble_BestOfFamily” and “StackEnsemble_AllOfFamily” as well as base statistical machine learning models DeepLearning, DRF, XRT, GBM, and GLM, here define AUC is defined as a stopping metric. Out of that “StackEnsemble_BestOfFamily” was giving an AUC of 71.08% and in base models, DRF was giving the highest Accuracy 88.02%, and an AUCof 70.5%. While in the interpretation SHAP, PDP, and ICE techniques gave a very good explanation for every individual result as well as a prediction of the whole data set.

Hence with the help of AutoML techniques, multiple machine learning models have been created in a very short time without wasting much time on data preparation, Data exploration, Feature engineering, model selection, model training, and hyperparameter tuning. With the help of SHAP, PDP, and ICE any individual result could be explained to the customer or the end-user.

Keywords— ***AutoML, metalearner, StackEnsemble_BestOfFamily, StackEnsemble_AllOfFamily, DeepLearning, DRF, XRT, GBM, GLM, AUC, SHAP, PDP, and ICE.***

I. INTRODUCTION (*HEADING 1*)

Credit risk is the likelihood that the borrower won't be able to pay back his debt on time or in whole. It refers to the possibility that the lender won't get the principal loaned or the interest due in a timely manner. This has the result of interfering with creditors' cash flow and raising the cost of collection. In dire circumstances, it can be necessary to write off all or a portion of the debt, incurring a loss for the lender. Knowing with absolute certainty how likely someone is to fail on a debt is very tough and complex. A competent evaluation of credit risk can also lessen the chance of losses due to default and late payments. The lender receives interest payments from the borrower as compensation for taking on the credit risk. Lenders or investors will either charge a higher interest rate or decline the loan opportunity entirely if the credit risk is higher. For instance, an applicant for the same loan with excellent credit and consistent income will pay a lower interest rate than one with a poor credit history [1].

A person's credit risk is influenced by a wide range of variables. As a result, determining the borrower's credit risk is an extremely challenging undertaking. Credit risk modelling has become important since so much money depends on our ability to predict a borrower's credit risk accurately. Credit risk modelling is the practise of utilising data models to ascertain two crucial facts. The likelihood that the borrower will miss a loan payment is the first. The second is how this default will affect the lender's financial situation [2].

To assess the credit risk of potential borrowers, financial institutions use credit risk models. Based on the validation of the credit risk model, they determine whether or not to authorise the loan and the loan's interest rate. With the development of technology, new approaches to modelling credit risk have appeared, such as credit risk modelling in R and Python. These include modelling credit risk utilising the newest analytics and big data tools. The way credit risk is modelled has also been impacted by other variables, such as the growth of economies and the ensuing rise of various types of credit risk[3].

Financial organisations have created sophisticated methods for calculating and controlling credit risk across all product categories of companies. A good understanding of frequently employed methods would, in the eyes of a regulator, increase the oversight of financial institutions. The necessity to calculate the amount of capital required to support the bank's exposures initially sparked interest in credit risk models. The job of the credit risk model is to take general economic conditions and specific variables as inputs and produce a credit spread as an output. There are two primary categories of credit risk models in this regard: structural and scale models. Based on the value of a company's assets and liabilities, structural models are used to determine the likelihood that the company would fail. If a company's assets are worth less than the amount of debt that must be repaid, it goes into default. Reduced models consider a random, exogenous source of failure[4].

With the help of the credit risk modelling methodology, risk measurement and management may be done in a flexible and personalised way. Models respond to changes in business lines, credit quality, market factors, and the economic climate as a result of their design and are influenced by these changes [5].

Additionally, the models reflect concentration risk within the portfolio and give banks the ability to examine marginal and absolute risk contributions. These model characteristics can help the bank's overall credit culture. The degree to which credit management has embraced models. Between banks, there are significant differences in how economic capital is distributed. While some banks have put in place systems to capture the majority of exposures across the organisation, only others do so for a specific business line or legal entity. In addition, banks frequently create distinct models for corporate and retail exposures, and not all banks account for both kinds of exposures. Internal model applications range from straightforward to straightforward and sophisticate [6].

Only a tiny portion of the banks the working group investigated already use the results of credit risk models in active portfolio management, but a sizable portion stated they intended to do so in the future. Applications now used include:

a)Calculating exposure and concentration limits;
b) Defining holding objectives for syndicated loans;
c) Pricing based on risk;
d) Enhancing the portfolio's risk and return profile;
e) Evaluation of business lines' or managers' risk-adjusted performance using risk-adjusted return on equity ("RAROC")
f) Capital allocation in the economy. Institutions also use model estimates for direct computations or for validation purposes to generate or confirm reserves for loan losses [7].

The Working Group acknowledges that improved internal risk management in banking organisations can be a result of credit risk modelling. However, before employing the models in the process of establishing regulatory capital needs, the key challenges related to data limitations and model validation must be overcome. Description of the default process and other elements that affect loans The absence of previous credit and loan performance data as well as other modelled variables greatly restricts quality. Since credit risk is measured across lengthy time horizons, which means the data span many years, the specification issues become more severe. To accurately estimate important factors, several credit cycles can be necessary. As a result of the present constraints, model parameters frequently involve some degree of simplification of assumptions and information collection from various sources. Since doing sensitivity testing of the model's vulnerability to such assumptions is not currently conventional practise, the impact of these alternatives on model risk estimation is uncertain[8] .

II. LITERATURE REVIEW

In a commercial bank, credit management is a challenging functional area. It requires skilled handling, accurate risk assessment at each level, and sufficient assurance of the security of monies exposed. Despite greatest attempts, it is unable to create solid security requirements, which causes credit to be unstable or unpaid on a frequent basis. Therefore, effective risk management, asset and liability management techniques, and continual search for safer criteria for risk elimination constitute credit management. Such risk management expertise has been developed and put into use to help mitigate risk rather than eradicate it[9] .

Credit risk is typically understood as the risk of default, or the risk of losing money if the borrower or counterparty fails to pay the bank the amount owed (principal or interest) according to a pre-arranged repayment schedule, on time. Value risk, or the danger of losing value as a result of the borrower's shift to a lower credit rating (opportunity costs connected with incorrectly pricing a new loan risk level), would be added to the definition to make it more complete[10] .

Banks have devised procedures that enable them to quantify these risks and so derive the necessary amount of capital to maintain their company, known as economic capital, in order to hedge against volatility in default/impairment levels (as well as other types of risk). The method for determining the required minimum regulatory capital is outlined in Pillar 1. Basel-I stated that this computation only took into account credit risk; however, in 1996, a calculation for market risk was introduced. Basel II increases the price to reflect operational risk.

A lot of effort needs to be done at the bank level as the entire financial sector worldwide works to implement the 2004 directive-based II Accord in some shape and intensity. Credit Risk Management gives users the tools they need to develop the necessary management structure, rules, processes, and practises for credit risk measurement[11].

One of Basel II's main advances is that it gives lenders an option between I and II, in contrast to Basel I, which only provided a single method for computing regulatory capital for credit risk. Using a standardised approach, Basel I is expanded upon by classifying exposures into several risk categories. Nevertheless, each risk category historically had a defined risk weight, e.g. The risk weights for the three Basel II categories (loans to governments, businesses, and banks) are established by the borrowers' given external credit ratings. Loans backed by residential real estate, among other categories that continue to be subject to fixed risk weights under Basel II, will bear a risk weight of 35%, up from 50% previously, if Loan-To-Value (LTV) climbs to 80%. This reduced weighting is in acknowledgement of historically low loss rates that frequently originated in residential mortgage loan portfolios across a wide range of economic situations in many nations [12].

2. A Basic approach Internal Rating Based (IRB) - The IRB approach will enable lenders to create own models for calculating their regulatory capital need. Lenders assess the Probability of Default (PD) under the foundation IRB approach, and supervision sets the numbers for Loss Given Default (LGD), Exposure At Default (EAD), and Maturity Exposure (ME). For each exposure or kind of exposure, these numbers are fed into the lender's appropriate risk weight function to produce risk weights[13] .

Under this strategy, banks can utilise credit risk limiters (collateral, guarantees, and credit derivatives) to lower capital depending on the market risk of the collateral instrument that others hold and the external credit assessment of recognised guarantors. Residential mortgages, small and medium-sized enterprises (SME), and retail exposures all had their risk weights reduced. This method significantly differentiates exposures from transactions in order to increase the resulting risk-sensitive capital ratios[14] .

III. RESEARCH METHODOLOGY

In case of Machine learning methodology, Data related steps same as in traditional approach but in place of manually understanding the relationship between data trends and their outcomes different Machine learning algorithms has been used which gives good in predictions based on different inputs. But the Machine learning algorithms are tedious and time consuming and it difficult for non- technical person to develop a statistical model which can give good accuracy.

In this project in place of traditional methodology, AutoML has been used. With the use of AutoML most of the Data related steps are same as traditional approach but for the selecting different machine learning algorithms, training and hyperparameter tuning could be automated.

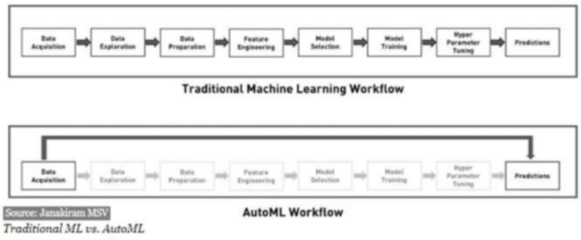

Figure 3.1 - AutoML vs Traditional Machine learning workflow

In Figure 3.1 it has been shown that in the case of AutoML basic time-consuming steps of data modeling like Data explorations, Data Preparation, Feature Engineering, Model selection, Model training, and Hyperparameter tuning can be bypassed. So in this way, a lot of time could be saved that researchers generally waste during these processes and concentrated on Data Collection and deployments of the best model[15].

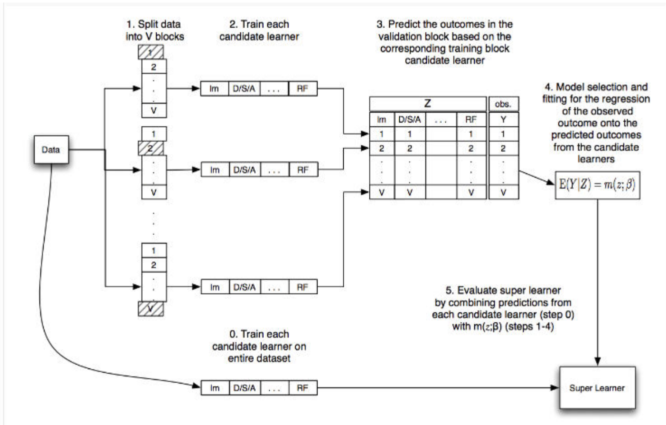

Figure 3.2: Process flow in AutoML

In Figure 3.2 Super learners can be described as the stack of different ensemble models, In this process, certain steps have been followed:

1. The cross-validation has been selected for the training dataset.
2. The maximum number of base models has also been specified.
3. For every base model these set has been followed:
   3.1. Base models have been evaluated by using cross-validation
   3.2. Store their predictions
4. Then trained the base models on the training dataset and stored predictions.
5. Meta-learners on the stored predictions
6. The prediction has been done on the Test Data set[16].

Figure 3.3: AutoMLWorkFlow

In Figure 3.3 steps followed by AutoML have been shown, they are as follows:

1. Load and train and Test data.
2. Specify the response feature.
3. Running the AutoML by specifying the stopping criteria e.g. Maximum number of models, the maximum time to train is stopping metric or stopping rounds.
4. Analyze different models on the leader board and explore them based on our requirements usually ensemble stack model best of the family or All models give better accuracy.
5. The model could be saved and deployed in production for further uses[17].

IV. DATA MODELING

Data modelling in software engineering is the practise of utilising formal approaches to simplify a diagram or data model of a software system. It involves communicating information and data using text and symbols. A data model offers a design framework for creating a new database or redesigning old applications.

model_id	auc	logloss	aucpr	mean_per_class_error	rmse
StackedEnsemble_BestOfFamily_3_AutoML_8_20220716_204454	0.710815	0.326445	0.286215	0.351219	0.306801
StackedEnsemble_AllModels_2_AutoML_8_20220716_204454	0.710589	0.326455	0.286061	0.344115	0.306765
StackedEnsemble_AllModels_1_AutoML_8_20220716_204454	0.710347	0.326806	0.284202	0.351165	0.306946
StackedEnsemble_BestOfFamily_2_AutoML_8_20220716_204454	0.710149	0.326896	0.284905	0.346986	0.30696
DRF_1_AutoML_8_20220716_204454	0.704951	0.33437	0.286943	0.34251	0.309266
GBM_2_AutoML_8_20220716_204454	0.701234	0.331476	0.263136	0.352149	0.309638
XRT_1_AutoML_8_20220716_204454	0.700893	0.337806	0.271597	0.354919	0.312968
GBM_4_AutoML_8_20220716_204454	0.700447	0.334691	0.270076	0.351465	0.309524
GBM_3_AutoML_8_20220716_204454	0.700335	0.332374	0.260778	0.353529	0.309812
StackedEnsemble_BestOfFamily_1_AutoML_8_20220716_204454	0.699799	0.33095	0.261124	0.358147	0.30917
GBM_1_AutoML_8_20220716_204454	0.69844	0.332399	0.261681	0.357177	0.309774
GBM_5_AutoML_8_20220716_204454	0.696926	0.332226	0.259248	0.364794	0.309984
GLM_1_AutoML_8_20220716_204454	0.668986	0.340973	0.218211	0.369995	0.313912
DeepLearning_1_AutoML_8_20220716_204454	0.612563	0.397489	0.195069	0.415548	0.32298

Figure 4.1: Leader Board of AutoML

In Figure 4.1, the Leader Board of the AutoML has been shown, where 14 machine learning models with different metrics have been shown, out of that Best of Family ensemble machine learning models outperformed other ensembles as well as base models.

V. DATA EVALUTAION

Plans for evaluation should outline the methods and sources used to gather data. Both quantitative and qualitative data has to be gathered within a structure that is in keeping with programme goals, stakeholder expectations, and project schedules.

```
metalearner.varimp()
```

```
[('DRF_1_AutoML_8_20220716_204454',
  0.45089760422706604,
  1.0,
  0.5377771520927915),
 ('XRT_1_AutoML_8_20220716_204454',
  0.21615146100521088,
  0.4793803714609224,
  0.25779981093343934),
 ('GBM_2_AutoML_8_20220716_204454',
  0.10483089834451675,
  0.23249380205561107,
  0.12502985474869172),
 ('DeepLearning_1_AutoML_8_20220716_204454',
  0.03858282417058945,
  0.08556892697784141,
  0.046017013857779586),
 ('GLM_1_AutoML_8_20220716_204454',
  0.027984146028757095,
  0.06206319520532349б,
  0.03337616836729786)]
```

Figure 5.1: Importance of different Base models in Best of Family Stacked Model.

In Figure 5.1, it was shown how the different base models contribute to the best family stacked ensemble model. Here it has been noticed the DRF was the playing most important role in making the decision.

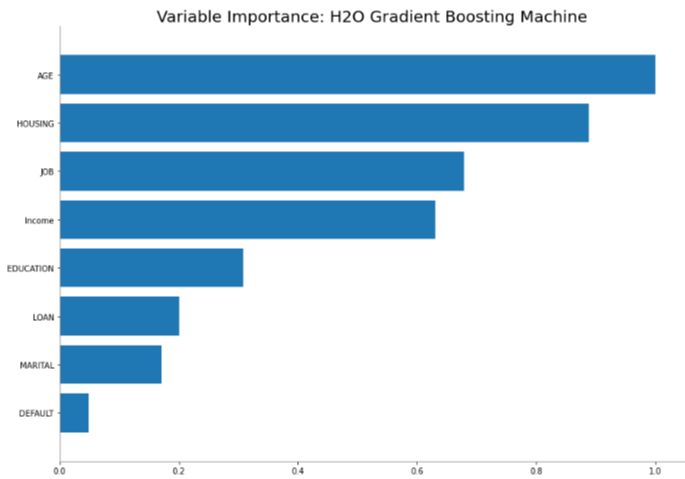

Figure 5.2: Variable Importance for the Gradient boosting model

In Figure 5.2, it is shown how the Gradient boosting base model treats to different independent features for making any decision. As Gradient boosting gave more importance to Age, housing and Job as compare to Marital, Loan and Education.

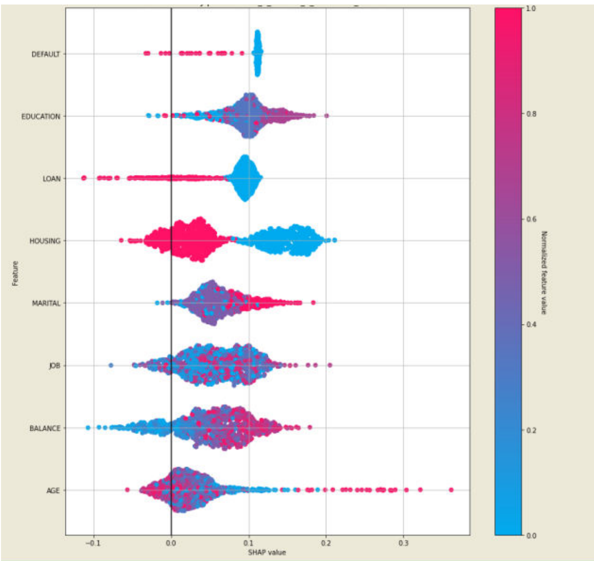

Figure 5.3: SHAP Explanation

In Figure 5.3, A positive SHAP number indicates a positive impact on prediction, which causes the model to predict 1, according to the SHAP explanation (e.g. Loan approved). The model predicts 0 since a negative SHAP score indicates a negative impact (e.g. Loan did not approve).

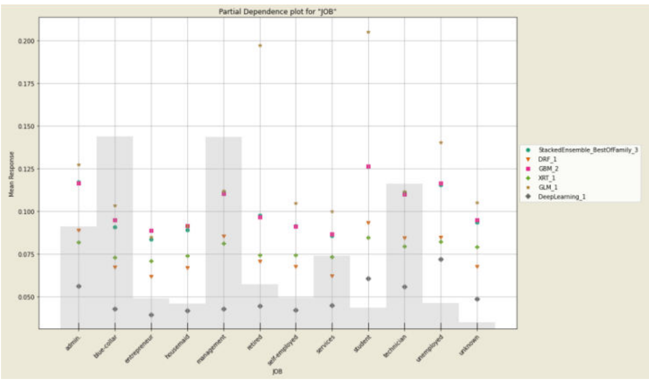

Figure 5.4: Partial Dependence Plot Effect of Job on different models

In Figure 5.4, it has been shown the effect of Job independent features on different ensembles as well as base models with the help of PDP. GLM_1 giving more importance Admin., self-Employed and unemployed person. Whereas GBM_2 gave the importance person who are working as Admin., management, student and unemployed.

Individual row prediction

AGE	JOB	MARITAL	EDUCATION	DEFAULT	HOUSING	LOAN	LOAN_STATUS	Income
42	admin.	married	secondary	no	yes	no	no	1173

predict	no	yes
no	0.934694	0.0653063

Figure 5.5: Single Row prediction by Best Base Model

In Figure 5.5, checking the individual customer result and their probability of approval and not approval with the best base model. In this case as per base model probability of rejecting the loan application is 93.47% and for getting approval is only 6.53%. Which is same as real Loan status.

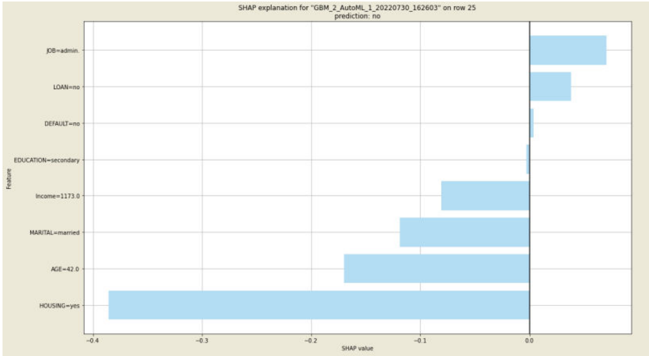

Figure 5.6: SHAP Explanation of the result

Figure 5.6, it has shown the explanation of individual customer results with the help of SHAP. Job and loan positively impact the chance of loan approval whereas Housing, age, marital status, and income effect are more in rejecting their loan request.

VI. ANALYS AND RESULT

In this research, it has been shown that with the use of Automate Machine learning technique in the combination of different explanations e.g. SHAP, PDP and ICE. Different complex, as well as base models, can be developed in a short time with minimal knowledge of programming and compared with different metrics. Researchers could save a lot of time in developing, training, or tuning the different machine learning models, they could spend that time on data collection and understanding it. On rejecting any loan application end user can explain the reason or features behind that so the customer can also be satisfied with the explanation.

VII. CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

So it has been shown Automated Machine learning can be a very useful technique in the field of financial and banking sectors. Researchers or machine learning engineers can use Automated Machine learning in a combination of explanation techniques in providing very efficient models to the industry and provide insight also how the complex model is working and explain every prediction done by it.

VIII. REFERENCES

[1] "Credit Risk: Definition, Role of Ratings, and Examples." https://www.investopedia.com/terms/c/creditrisk.asp (accessed Sep. 14, 2022).

[2] C. Bluhm, L. Overbeck, and C. Wagner,

"Introduction to Credit Risk Modeling," *Introd. to Credit Risk Model.*, Apr. 2016, doi: 10.1201/9781584889939.

[3] "Credit Risk: Modeling, Valuation and Hedging - Tomasz R. Bielecki, Marek Rutkowski - Google Books." https://books.google.co.in/books?hl=en&lr=&id=GWzmCAAAQBAJ&oi=fnd&pg=PR5&dq=Credit+Risk:+Modeling,+Valuation+and+Hedging,+Tomasz+R.+Bielecki,+Marek+Rutkowski&ots=6Zobowx0pC&sig=StBEDSuUxXWSh9ibAuqeFcmUF6I#v=onepage&q=Credit Risk%3A Modeling%2C Valuation and Hedging%2C Tomasz R. Bielecki%2C Marek Rutkowski&f=false (accessed Sep. 14, 2022).

[4] C. A. Fernandes and C. Dissertação, "Structural Models in Credit Risk," 2014.

[5] D. J. Teece, "Business models and dynamic capabilities," *Long Range Plann.*, vol. 51, no. 1, pp. 40–49, Feb. 2018, doi: 10.1016/J.LRP.2017.06.007.

[6] "RBC20 - Calculation of minimum risk-based capital requirements." https://www.bis.org/basel_framework/chapter/RBC/20.htm?inforce=20230101&published=20201126 (accessed Sep. 14, 2022).

[7] "Risk-Adjusted Return On Capital (RAROC)." https://www.investopedia.com/terms/r/raroc.asp (accessed Sep. 14, 2022).

[8] J. Bessis, "Risk Management in Banking," p. 840 p., 2009, Accessed: Sep. 14, 2022. [Online]. Available: https://hal-hec.archives-ouvertes.fr/hal-00494876

[9] "Reserve Bank of India - Index To RBI Circulars." https://www.rbi.org.in/scripts/BS_CircularIndexDisplay.aspx?Id=4964 (accessed Sep. 14, 2022).

[10] "Amendment to the capital accord to incorporate market risks," Nov. 2005, Accessed: Sep. 14, 2022. [Online]. Available: https://www.bis.org/publ/bcbs119.htm

[11] "Basel II: Revised international capital framework." https://www.bis.org/publ/bcbsca.htm (accessed Sep. 14, 2022).

[12] B. - Bank for International Settlements, "Y V Reddy: Challenges and implications of Basel II for Asia (Central Bank Articles and Speeches)," 2006.

[13] "Modelling credit risk | Bank of England." https://www.bankofengland.co.uk/ccbs/modelling-credit-risk (accessed Sep. 14, 2022).

[14] "Zhang, Xuan (2017) Essays in credit risk management. Abstract Credit risk management is becoming more and more important in recent years. Credit risk refers to the risk that an obligor." https://dokumen.tips/documents/zhang-xuan-2017-essays-in-credit-risk-management-abstract-credit-risk-management.html?page=1 (accessed Sep. 14, 2022).

[15] "Using an AutoML H2O Model to Predict Attrition and LIME to Explain the Predicted Class - DZone AI." https://dzone.com/articles/using-an-automl-h2o-model-to-predict-attrition-and (accessed Sep. 14, 2022).

[16] "Super Learning | SpringerLink." https://link.springer.com/chapter/10.1007/978-1-4419-9782-1_3 (accessed Sep. 14, 2022).

[17] "A Deep dive into H2O's AutoML. An overview of Automated Machine… | by Parul Pandey | Towards Data Science." https://towardsdatascience.com/a-deep-dive-into-h2os-automl-4b1fe51d3f3e (accessed Sep. 14, 2022).