25th January 2023

AI and Machine Learning for Credit Rating Models – Part III

The power of machine learning

fintegral

AI and Machine Learning for Credit Rating Models

Overcoming traditional modelling challenges with ML approaches

Many of the traditional modelling challenges facing firms today can be tackled with machine learning (ML) approaches. Such techniques include reframing the objectives and problems, training ML algorithms on the data to learn more about their attributes and/or by adopting an ensemble approach marrying both traditional and ML approaches.

Modelling challenge	Possible solutions	Potential advantages and disadvantages
Use of unstructured data Traditional models typically lack the capabilities to consume unstructured data types.	• **Text and sentiment analysis:** Methods like Natural Language Processing (NLP) are currently being used to extract information from unstructured articles (e.g. for probability of default (PD) models).	• **Increased feature identification:** NLP can help identify new additional features for models and can be used to build early warning systems.
Understanding data and data quality Data cleaning and analysis of new data can be a tedious task demanding high and often manual efforts from data specialists. In addition, this task can become time consuming requiring revisiting during the project cycle if the initial data cleaning activities were not effective.	• **Outlier detection methods:** In addition to traditional outlier detection tools such as graphical methods, standardised residuals, and Cook's distance measure, ML approaches including clustering may support the data cleaning phase. • **ML training:** The training of ML models on the data followed by the analysis of these models may lead to an improved understanding of the data's inner structure, data dependencies and their importance to the final model's target variable.	• **Greater efficiency:** Better handling of data quality issues and outliers lead to more efficient models. • **Enhanced feature engineering:** Deeper analysis of the data structure may support new features being identified leading to improved model performance. • **Mutual approaches:** By using ML techniques in the data cleaning stage combined with traditional approaches for the final model, will allow developers to harness the benefits of both techniques whilst upholding model explainability.
Managing multi-sector portfolios The segmentation of corporate, small and medium-sized enterprises (SME) and/or non-retail portfolios traditionally involves a multitude of (sometime iterative) manual analyses, making this typically a very time-consuming task.	• **Clustering methods:** K-Means and K-Nearest Neighbours (KNN) may be used to combine and segment sectors by assessing the correlations between the features and uncovering hidden patterns.	• **Improved segmentation:** K-Means clustering can dissect large datasets into smaller datasets with similar characteristics. • **Increased simplicity:** Clustering is simple, flexible and interpretable. • **Increased manual intervention:** The optimal number of clusters must be specified manually.
Selecting features and managing multicollinearity During the feature selection phase, models may produce an excessive number of highly correlated features leading to models being inefficient. These issues are exacerbated when dealing with Big Data.	• **Principle Component Analysis (PCA):** PCA aims to reduce the dimensions of datasets while preserving the maximum amount of information achieved by aggregating features into new features. • **Regularisation:** Least Absolute Shrinkage and Selection Operator (LASSO) aims to reduce the magnitude of coefficients, sometimes to zero, by applying penalty terms.	• **Reduced overfitting:** PCA and LASSO can help reduce model overfitting in high-dimensional datasets caused by excessive features. • **Reduced variability:** LASSO reduces the statistical variability of high-dimensional data problems. • **Faster:** Fewer features allow ML algorithms to train and learn more quickly and efficiently.

AI and Machine Learning for Credit Rating Models

Managing ML related challenges

ML approaches have their own challenges with ensuring equality and managing the “black-box” phenomenon being key areas of concern. Additional testing and utilising specialised visualisation tools such as SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) will help to establish wider stakeholder acceptance including from regulators.

Modelling challenge	Possible solutions	Potential advantages and disadvantages
Improving equality and fairness The detection or avoidance of unfair/discriminatory behaviour is made more difficult in ML models than traditional models due to their “black-box” nature. Regulators and local laws often restrict the use of “protected attributes” in models.	• **Input testing:** Proper testing of input data for possible biases (using statistical testing, but more importantly qualitative assessment on data sourcing[1]). • **Model testing:** Testing of the fairness of the model for certain demographic populations[2]. • **Model interpretability/explainability:** Check for fairness by Subject Matter Experts (SMEs) using model interpretability approaches.	• **Greater understanding and acceptance:** Proper testing of models against unfair/discriminatory behaviour greatly improves the acceptance by regulators and stakeholders.

Mitigating the “black-box” phenomenon
ML models are often complex and cannot be easily interpreted, even by experts. This black-box effect remains a key challenge in the adoption of ML, as both regulators and stakeholders require a strong understanding of the models. Additional guidance on how to improve model interpretability is required.
	• **Model design:** The design of the models play an intricate part in ensuring models are explainable. Model developers can choose to either design models from the ground up whilst upholding model explainability; or use external interpretability models (global or localised) which will analyse the main model to gain a stronger understanding of the predictions.
	  • Global interpretability models focus on the overall decision process and cannot explain the cause of each prediction.
	  • Localised interpretability models focus on individual predictions and the driving forces behind each prediction by using popular methods such as SHAP and LIME.
		• **Enhanced visualisation:** SHAP and LIME can visualise the importance and influence of features on the predicted values.
		• **Increased bias:** Localised models like Partial Dependency Plots (PDP), Accumulated Local Effects (ALE), and Individual Conditional Expectation (ICE) may be subject to bias when variables are highly correlated.
		• **Increased regulatory guidance required:** Clarity from regulators is required to help firms understand when ML-based models are sufficiently explainable and have met regulatory expectations.

[1] Historical data sets from different eras with different social values tend to contain discrimination against certain demographic groups. This discrimination is often “learned” by ML models.
[2] Simpson’s Paradox: For certain data sets, relationships may only be visible when data is analysed in sub-groups rather than being combined. Otherwise, this may lead to “spurious discrimination”, e.g., the “UC Berkeley gender bias”.

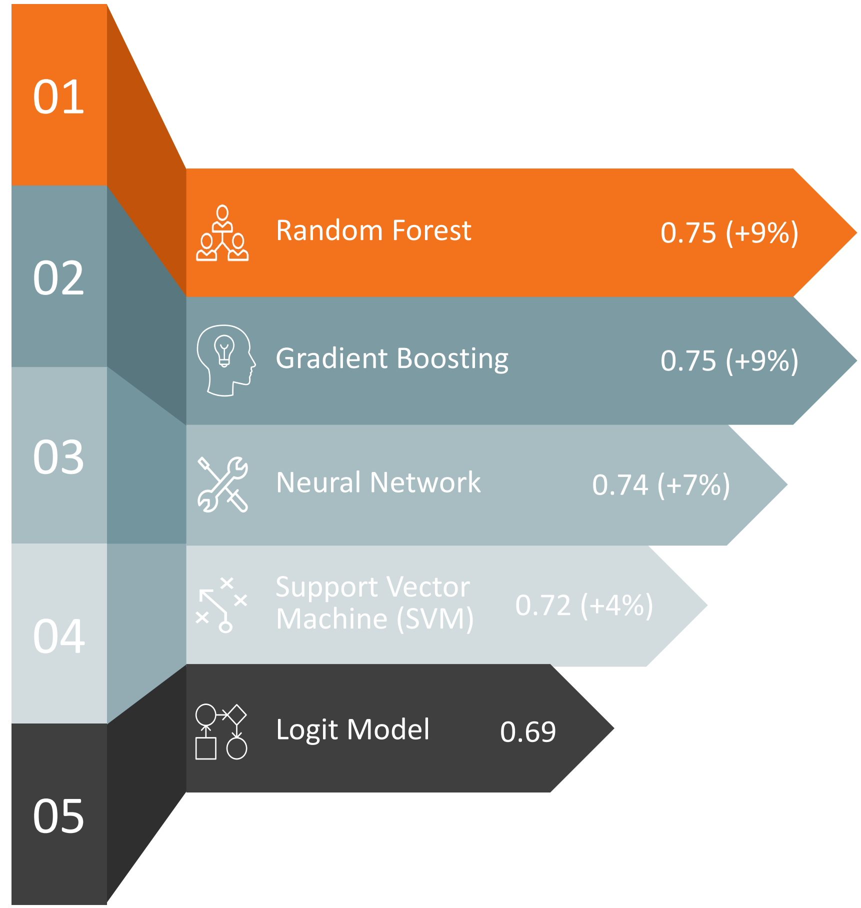
AI and Machine Learning for Credit Rating Models

Comparing traditional and ML algorithms

To demonstrate the power of ML algorithms, Fintegral conducted an analysis comparing the performance of different algorithms for modelling PDs on a real-world portfolio[1]. The Out-of-Sample (OOS) AUROC[2] scores illustrated all tested ML algorithms outperformed the traditional logit model, with the random forest and gradient boosting methods equally scoring the highest.

Technical comparison between traditional and ML algorithms

- To measure model accuracy and performance, Receiver Operator Characteristic (ROC) curves are used which plot the true positive rate (TPR) against the false positive rate (FPR).
- $TPR = \frac{TP}{TP+FN} \quad FPR = \frac{FP}{FP+TN}$
- The computation of TPR and FPR rates is a classical binary test designed to assess model sensitivity and specificity. In order to calculate the TPR and FPR rates, the True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN) values are required.
  - True Positive (TP) – Defaulted account correctly identified as being in default.
  - False Positive (FP) – Non-defaulted account incorrectly identified as being in default.
  - True Negative (TN) – Non-Defaulted account correctly identified as being not in default.
  - False Negative (FN) – Defaulted account incorrectly identified as being not in default.
- ROC curves are summarised by AUROC scores ranging between 0 and 1, with the higher the score, the better the model.
- Random forest and gradient boosting prove to have the highest AUROC scores at 0.75 representing a 9% favourable increase above the traditional logistic (logit) regression model's AUROC score of 0.69.
- Random forests are particularly attractive because of their less susceptible nature to overfitting.
- The logit model performs the worst compared to the other approaches with fewer accurate predictions (lower true positives and true negatives).



[1] Our analysis was conducted on a European retail mortgage portfolio.
[2] Area Under the Receiver Operator Characteristics

AI and Machine Learning for Credit Rating Models

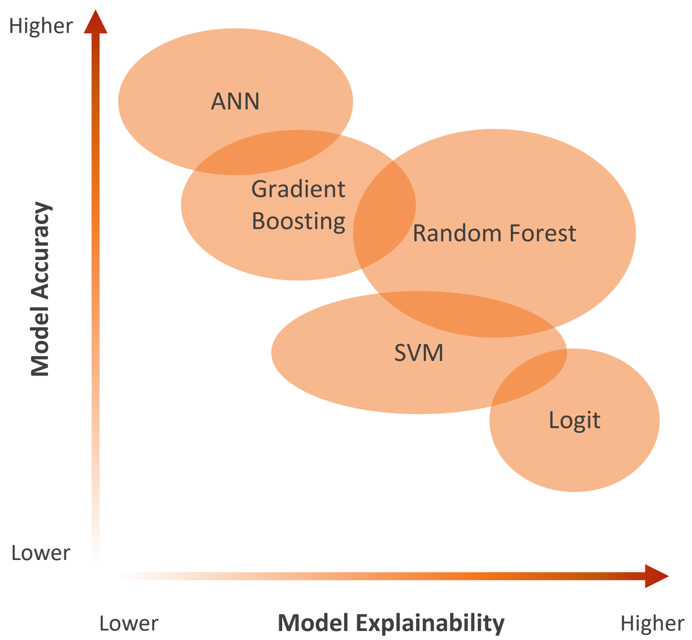
Conclusion and next steps

It is evident based on our assessment, ML algorithms for credit rating models have greater predictive capabilities, however, firms must balance the trade-off in model complexity and interpretability, as well as business costs. During the model development phase, it is crucial firms manage the risks associated with ML including model bias and discrimination.

Key takeaways

- **ML algorithms are powerful:** In our brief analysis on a European mortgage portfolio, random forest and gradient boosting outperformed the logit approach.
- **Balancing trade-off:** Despite the increase in model accuracy, there is an inherent trade-off in model explainability. An increase in model accuracy usually implies an increase in model complexity, however, this can be *partially* mitigated by new developments in explainable AI (XAI). Technically challenging models may require significant investment in additional resources and time spent, which will require consideration.
- **Simple yet complicated:** Some ML algorithms may be simple to understand at a holistic level but at a granular level, they can become complicated. For example, random forests are intrinsically straightforward but in real-world applications, they can expand to thousands of trees making them difficult to interpret. Firms should balance the trade-off between model accuracy and business costs as development of ML models may require additional resources.
- **Unique characteristics:** Despite the unique characteristics ML algorithms possess, the risks associated with the application of the model outputs remain on par with traditional methods.
- **Managing model risks:** Firms must emphasise on managing the risks of model bias and discrimination throughout the model development process. Moreover, focusing on data quality will help combat some of the risks plus aid in model interpretability.
- **No one size fits all solution:** By reframing the objectives and problems and breaking the challenges down into smaller tasks will allow developers to utilise the most optimal algorithm at each incremental stage, and harness the benefits of both traditional and ML approaches.

General Overview[1]



Next steps

In the next part(s) of our series on AI and ML for credit rating models, we will look at:

- Ways to optimise the hyperparameter space.
- Methods to overcome some of the risks associated with ML algorithms including looking at ways to develop trustworthy AI and implement XAI techniques.
- Aspects to consider when validating ML based models.

[1] This chart is not drawn to scale and the performance of the algorithms may differ significantly depending on the problem being tackled.

 • www.fintegral.com

Contact

Fintegral

London ¦ Frankfurt ¦ Zurich

www.fintegral.com

Dr. Andreas Peter
Managing Partner
Fintegral Group

+49 160 583 40 66
andreas.peter@fintegral.com

Fintegral Deutschland AG
Steinweg 5
60313 Frankfurt am Main
Germany

Dr. Tobias Kesselring
Senior Manager
Fintegral Schweiz AG

+41 79 271 19 00
tobias.kesselring@fintegral.com

Fintegral Schweiz AG
Claridenstrasse 35
8002 Zürich
Schweiz

Abdul Qaiyum
Senior Consultant
Fintegral UK Ltd.

+44 7496 363 298
abdul.qaiyum@fintegral.com

Fintegral UK Ltd.
City Tower, 40 Basinghall St.
London EC2V 5DE
United Kingdom