KU LEUVEN

UNIVERSITÀ DEGLI STUDI DI SALERNO

Boosting Credit Risk Data Quality using Machine Learning and eXplainable AI Techniques

The 8th Workshop on MIning DAta for financial applicationS, ECML-PKDD 2023, September 22, 2023 - Turin, Italy

Elena Tiukhova

Adriano Salcuni

Can Oguz

Marcella Niglio

Giuseppe Storti

Fabio Forte

Bart Baesens

Monique Snoeck

MIDAS

ECML PKDD 2023

Outline

- Problem statement
- Related work
- Feature engineering
- Anomaly detection
- Explaining anomalies with SHAP
  - SHAP for iForest
  - SHAP for Autoencoder
- Results
- Conclusions
- Future work

DQ framework

Motivation

Mary needs a loan

Bank evaluates Mary's financial situation

Good loan quality

Bank earns money => more capital for:

…compensating for losses

…granting new loans

Bad loan quality

Bank losses money => less capital for:

Motivation

Credit risk assessment
- Stability of financial system assurance
- Regulatory requirements compliance

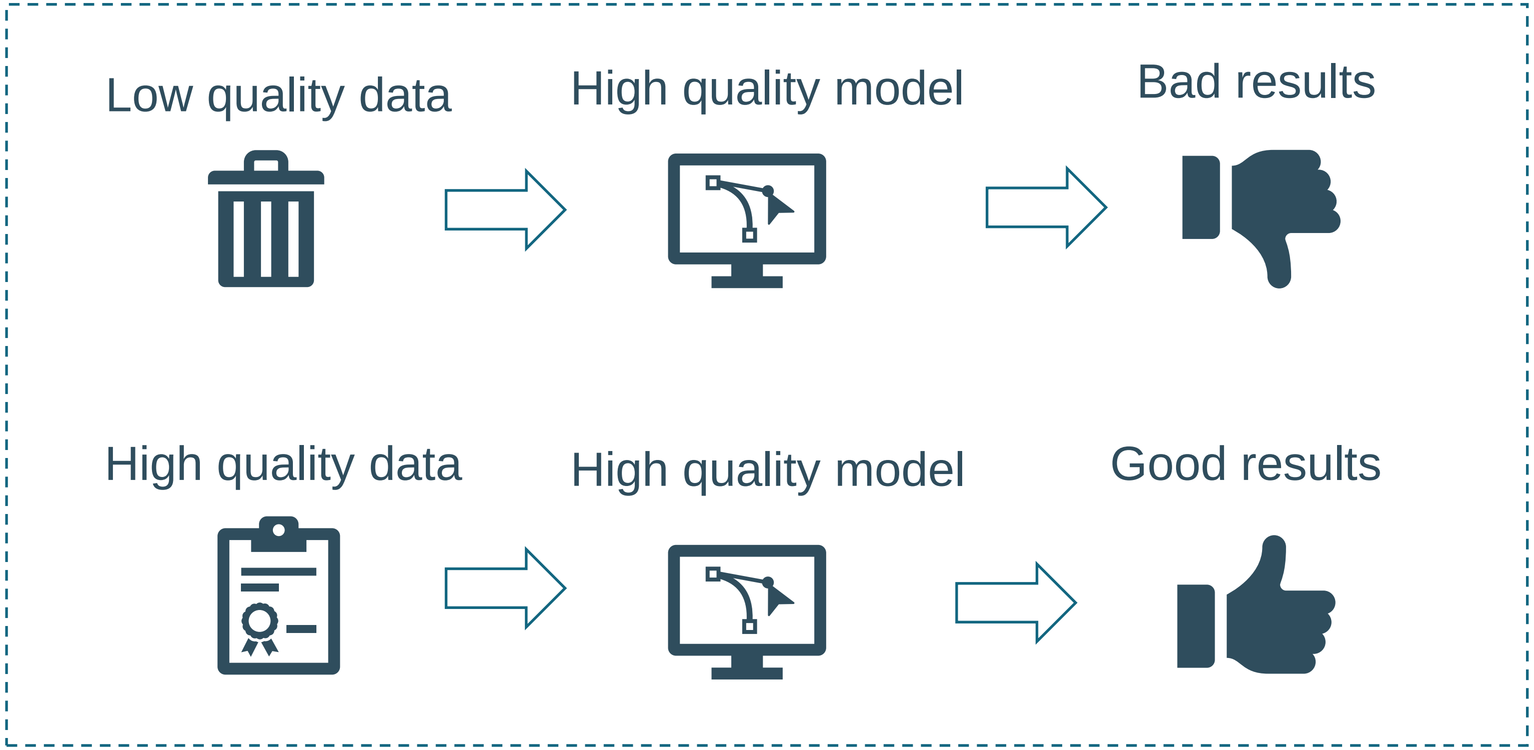
Risk-Weighted Assets (RWA) calculation
- Allocating a risk weight according to the estimated riskiness of various asset categories
- Keeping more capital to protect against potential losses on an asset with a greater risk weight

Credit risk data
- Decentralized with local systems being data entry points
- Consolidated in a central warehouse for reporting
- Diversity of data sources as a major reason for DQ issues (Moges et al., 2013)

Credit risk Data Quality assurance
- Manual evaluation not possible due to high volumes
- Automatic detection by Machine Learning (ML) models as a promising solution

Goal: investigate how ML models can be used to discover data quality (DQ) problems in credit risk data using a human-in-the-loop approach and bridge DQ and credit risk domains

DQ: Garbage in – garbage out



Related work

Credit risk

- Regulatory compliance
  - Basel III (BIS, 2017)
  - IFRS 9 (IFRS, 2014)

DQ in finance

- Data standardization and cleaning according to business rules and constraints (Nadinic & Kalpić, 2008)
- Total Data Quality Management (TDQM) for credit risk (Moges et al., 2013)
- ML for Anomaly detection in accounting entries (Bakumenko & Elragal, 2022)

eXplainable AI in Finance

- Auditing transparency (Zhang et al., 2022)
- Credit risk prediction explainability (Gramegna & Giudici, 2021; Bussmann et al., 2021; de Lange et al., 2022)

Credit risk overview

$$\text{Correlation (R)} = 0.12 \cdot \frac{\left(1-e^{-50 \cdot PD}\right)}{\left(1-e^{-50}\right)} + 0.24 \cdot \left(1 - \frac{\left(1-e^{-50 \cdot PD}\right)}{\left(1-e^{-50}\right)}\right)$$

$$\text{Maturity adjustment (b)} = \left[0.11852 - 0.05478 \cdot \ln(PD)\right]^2$$

$$\text{Capital requirement}^{13,14}\text{(K)} = \left[LGD \cdot N\left[\frac{G(PD)}{\sqrt{(1-R)}} + \sqrt{\frac{R}{1-R}} \cdot G(0.999)\right] - PD \cdot LGD\right] \cdot \frac{\left(1+(M-2.5) \cdot b\right)}{(1-1.5 \cdot b)}$$

$$\text{Risk-weighted assets (RWA)} = K \cdot 12.5 \cdot EAD$$

Source: Basel III

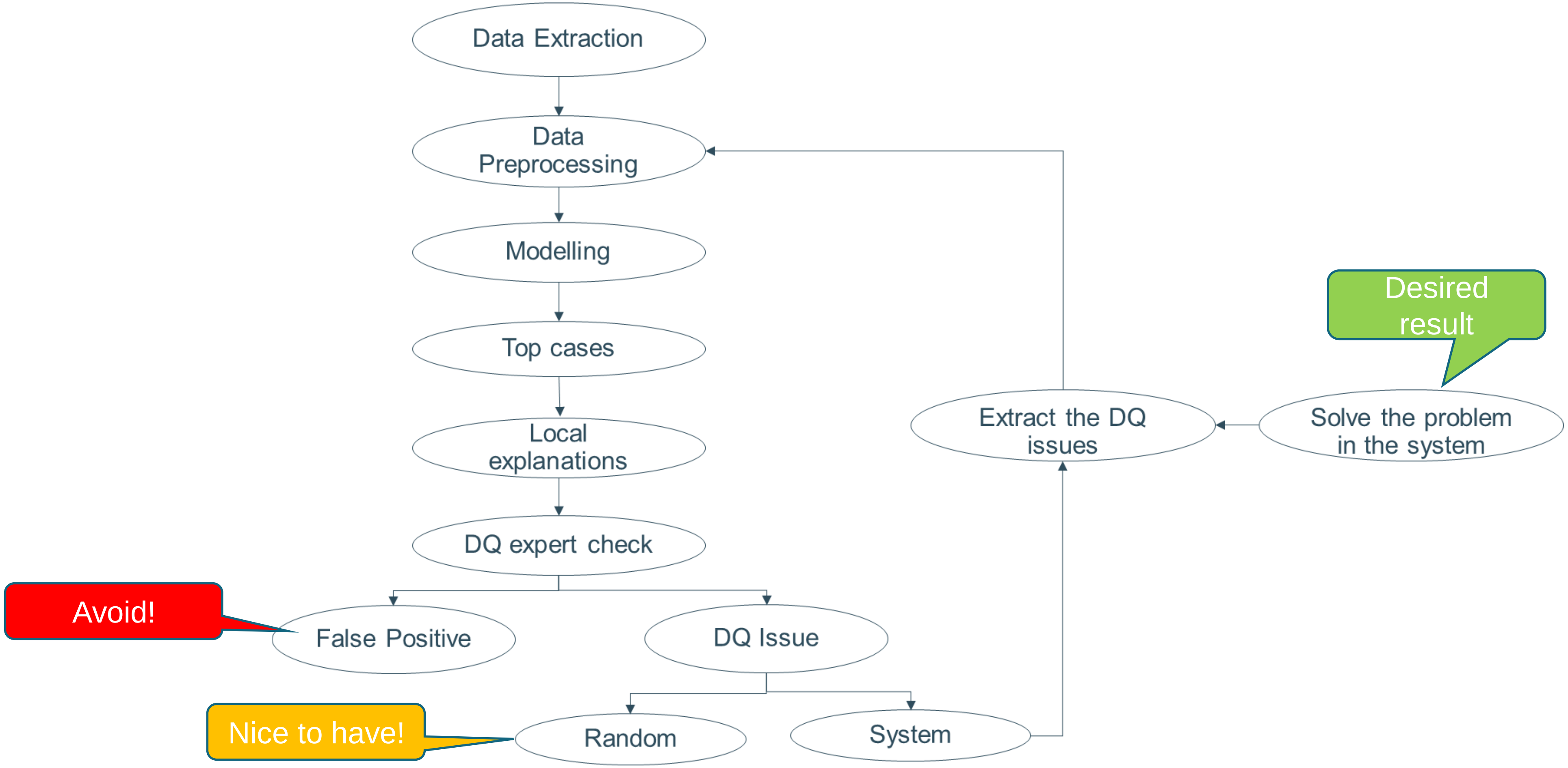
Credit risk estimation

- Internally determine the capital required to cover for unexpected losses
- Externally comply with accounting (IFRS 9) and banking (Basel III) standards

Risk-Weighted Assets (RWA)

- A standardized approach to measure credit risk
- Constrains the use of internal models
- Relies heavily on Probability of Default (PD), Loss Given Default (LGD) and Exposure At Default (EAD)

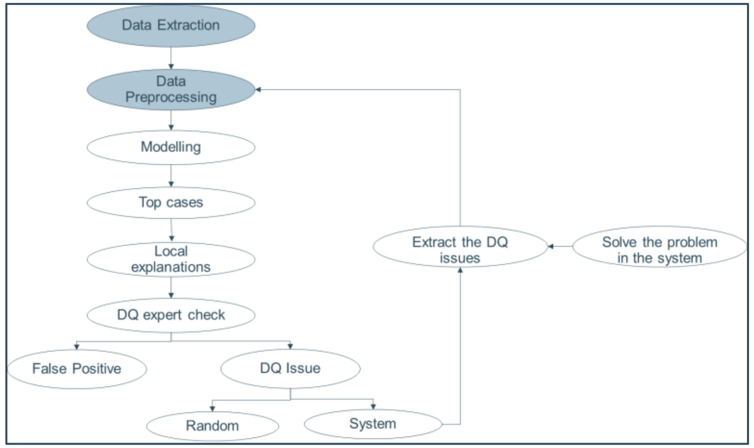
DQ framework



Feature engineering

~175k loans for SME & Mid. Corporates
from the multi-national bank

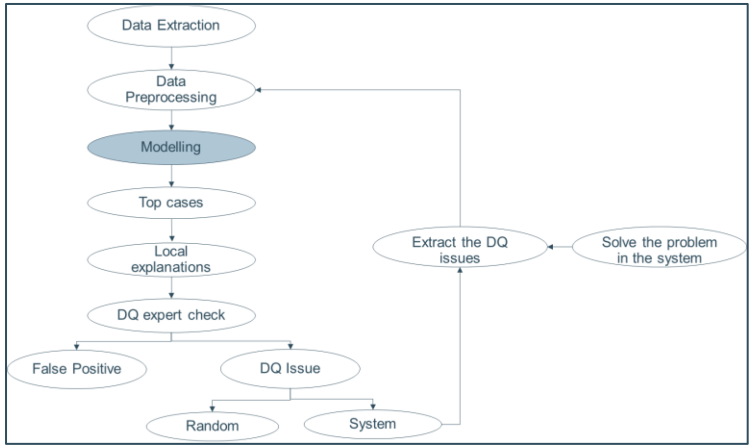
- Risk-Weighted Assets (RWA)
- Outstanding Amount (OS)
- Provisions (P)
- Maximal Limit (ML)
- Loss Given Default (LGD)
- Exposure-at-Default (EAD)
- Probability of Default (PD)



Ratio = $\ln \frac{V_t}{V_{t-1}}$: the way to incorporate changes in time

The **idea** is to check if the movements have business reasons or are anomalous.

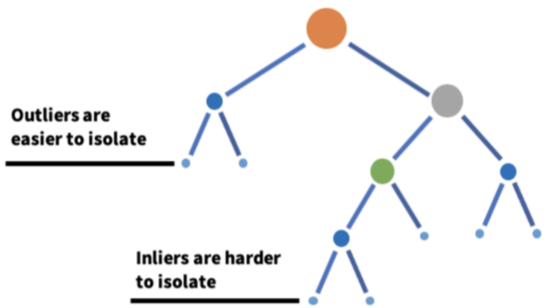
No DQ issues history → No labels → Unsupervised learning!

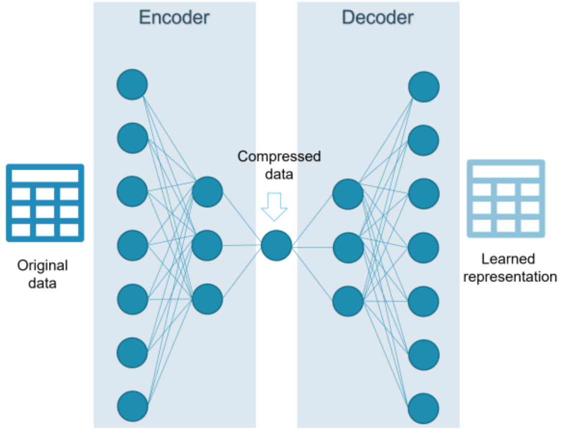


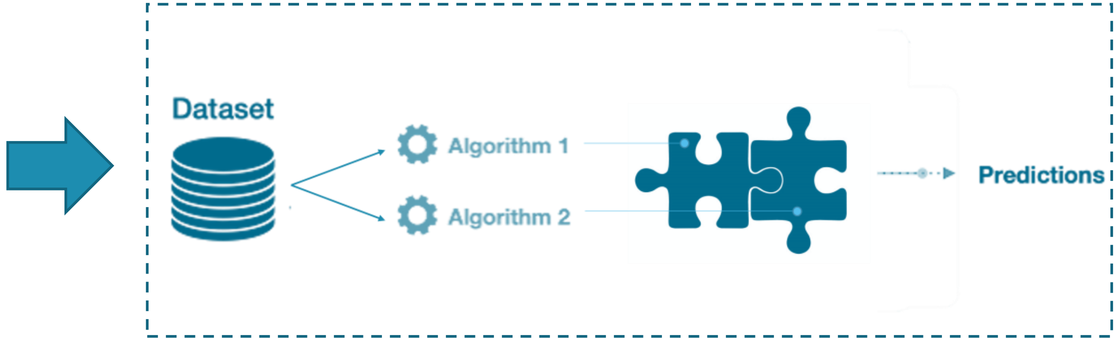
Isolation Forest   Autoencoder

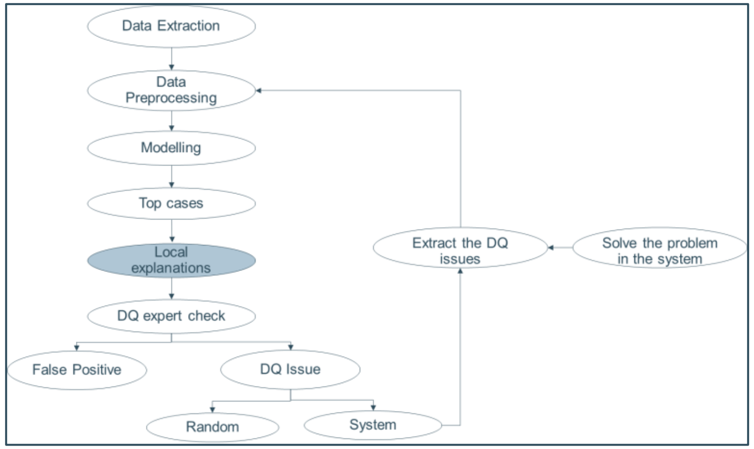
Why? - Among the best performing according to the literature and among the most adopted in practice (Han et al., 2022; Tiukhova et al., 2022)

Isolation Forest	Autoencoder
☐ Unsupervised outlier detection algorithm	☐ Unsupervised artificial neural network
☐ **based on the fact that outliers are "few and different"**, and therefore **easier to isolate**	☐ The encoder **compresses the input data into less dimensions** and **reduces the noise**
☐ Built using decision trees	☐ The **loss function is calculated to correct the reconstruction error** produced by the decoder









eXplainable AI: SHAP

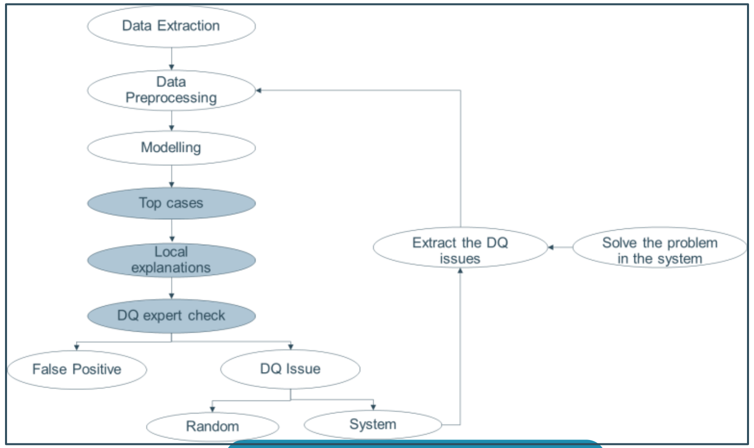
iForest

- SHAP TreeExplainer
- Explains the depth of the average path length of iTrees (Lundberg et al., 2020)
- Path length as a sum of additive feature contributions

Autoencoder

- SHAP KernelExplainer for Autoencoder
- Explains a reconstruction error: connection between the features with high reconstruction errors and the features affecting the reconstruction error the most (Antwarg et al., 2021)

DQ issues detection pipeline



Anomaly detection

- 2 models: iForest and Autoencoder
- Outlier scores transformati on: Min-Max Scaler

Scaled outlier scores →

Aggregation

- Currently: average of scaled outlier scores

Final outlier scores →

Scoring&XAI

- Top-20 cases to be checked by the DQ expert
- Local SHAP explanations: depends on the model's certainty

Top-K cases →

Feedback

- DQ issues
- False positives

Results of Round 2: top-20 cases

Loan	DQ-Issue?	Agg. Score	DQ issue features – feedback	SHAP top-3?
Loan 1	Yes	0,98	LGD	Yes
Loan 2	Yes	0,98	LGD & Max. Limit	Yes
Loan 3	No	0,97		
Loan 4	No	0,95		
Loan 5	No	0,92		
Loan 6	Yes	0,86	Relationship EAD, RWA, OS	Yes
Loan 7	Yes	0,86	Relationship EAD, RWA, OS	Yes
Loan 8	No	0,85		
Loan 9	Yes	0,85	EAD	Yes
Loan 10	Yes	0,79	Relationship EAD, RWA, OS	Yes
Loan 11	Yes	0,78	Relationship EAD, RWA, OS	Yes
Loan 12	Yes	0,78	Relationship EAD, RWA, OS	Yes
Loan 13	Yes	0,78	Relationship EAD, RWA, OS	Yes
Loan 14	Yes	0,78	Relationship EAD, RWA, OS	Yes
Loan 15	No	0,77		
Loan 16	Yes	0,75	Relationship EAD, RWA, OS	Yes
Loan 17	Yes	0,75	Relationship EAD, RWA, OS	Yes
Loan 18	No	0,75		
Loan 19	Yes	0,74	Relationship EAD, RWA, OS	Yes
Loan 20	Yes	0,72	OS	No

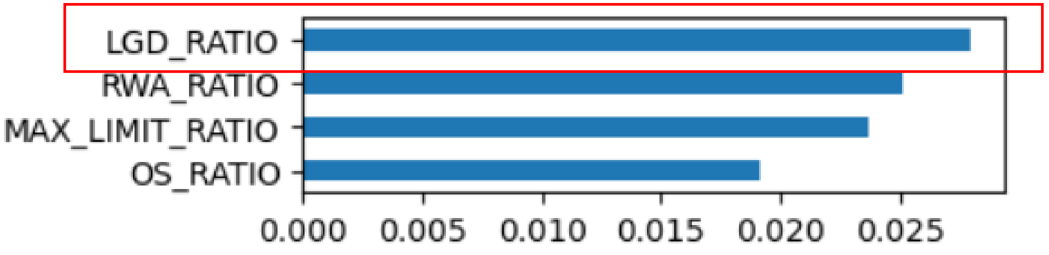
Potential systemic case!

Precision@20 = 70%

SHAP

Case 1: max. AE score

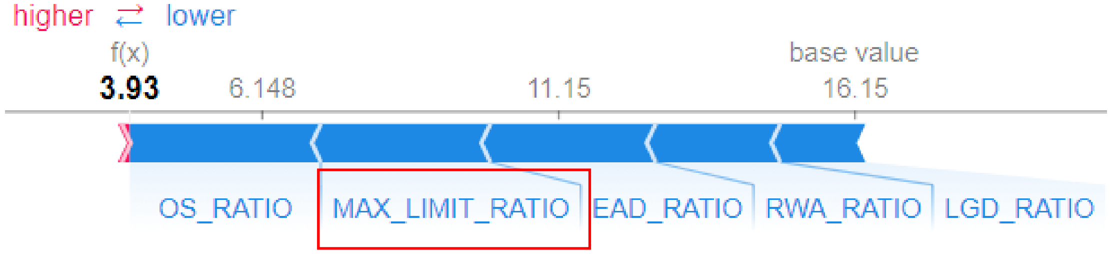
	RWA_ RATIO	OS_RATIO	PROVISION _RATIO	EAD_ RATIO	LGD_ RATIO	PD_RATIO	MAX_LIMIT_ RATIO
Loan 1	-16,84	-32,97	0	-17,76	0,99	0	-14,55



DQ expert feedback:
Risk class remained the same but LGD became higher

Case 2: max. IF score

	RWA_ RATIO	OS_RATIO	PROVISION_ RATIO	EAD_ RATIO	LGD_ RATIO	PD_RATIO	MAX_LIMIT_R ATIO
Loan 2	-16,96	-31,66	-28	-16,74	0,22	-1,27	-13,24



DQ expert feedback:
Max. Limit is set to 1 which is incorrect – potential systematic DQ issue

Conclusions

So far so good?

- Complementing static univariate DQ rules with dynamic multivariate ML model
- Novel ML DQ framework that can be generalized at other institutions
- RWA miscalculation analysis is still ongoing
- Transparent model output that increases trust in DQ ML techniques

Future work

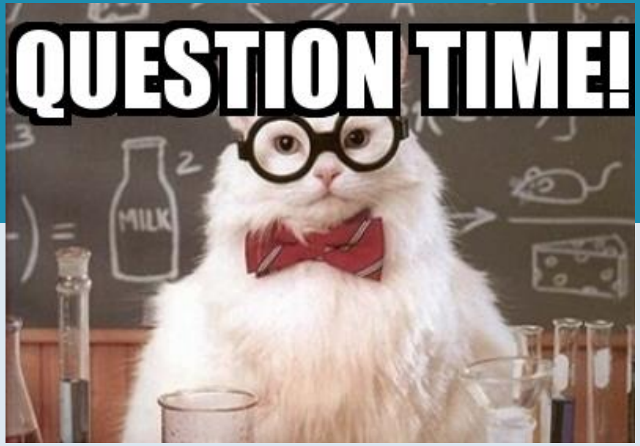
- DQ application for stakeholders
- LSTM Autoencoder for anomaly detection on daily time series data
- Model calibration based on the feedback

References

- Moges, H. T., Dejaeger, K., Lemahieu, W., & Baesens, B. (2013). A multidimensional analysis of data quality for credit risk management: New insights and challenges. *Information & management*, *50*(1), 43-58.
- Bank for International Settlements. (2017). Basel III: Finalising post-crisis reforms. Bank for International Settlements.
- IFRS - IFRS 9 Financial Instruments. (2014). https://www.ifrs.org/issued-standards/list-of-standards/ifrs-9-financial-instruments/
- Nadinić, B., & Kalpić, D.(2008). Data quality in finances and its impact on credit risk management and CRM integration. In Proceedings of the Third International Conference on Software and Data Technologies, pages 327-331 DOI: 10.5220/0001879103270331
- Bakumenko, A., & Elragal, A. (2022). Detecting anomalies in financial data using machine learning algorithms. *Systems*, *10*(5), 130.
- Zhang, C.A., Cho, S., Vasarhelyi, M.: Explainable artificial intelligence (XAI) in auditing. International Journal of Accounting Information Systems 46, 100572 (2022)
- Gramegna, A., Giudici, P.: Shap and lime: an evaluation of discriminative power in credit risk. Frontiers in Artificial Intelligence 4, 752558 (2021)
- Bussmann, N., Giudici, P., Marinelli, D., Papenbrock, J.: Explainable machine learning in credit risk management. Computational Economics 57, 203–216 (2021)
- de Lange, P.E., Melsom, B., Vennerød, C.B., Westgaard, S.: Explainable ai for credit assessment in banks. Journal of Risk and Financial Management 15(12), 556 (2022)
- Han, S., Hu, X., Huang, H., Jiang, M., Zhao, Y.: Adbench: Anomaly detection benchmark. Advances in Neural Information Processing Systems 35, 32142–32159 (2022)
- Tiukhova, E., Reusens, M., Baesens, B., & Snoeck, M. (2022, May). Benchmarking conventional outlier detection methods. In *Research Challenges in Information Science: 16th International Conference, RCIS 2022, Barcelona, Spain, May 17–20, 2022, Proceedings* (pp. 597-613). Cham: Springer International Publishing.
- Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.I.: From local explanations to global understanding with explainable ai for trees. Nature Machine Intelligence 2(1), 2522–5839 (2020)
- Antwarg, L., Miller, R. M., Shapira, B., & Rokach, L. (2021). Explaining anomalies detected by autoencoders using Shapley Additive Explanations. Expert systems with applications, 186, 115736.

KU LEUVEN

Thank you for your attention!



UNIVERSITÀ DEGLI STUDI
DI SALERNO