

Credit Risk in Fintech and Digital-Asset Markets: Integrating Machine Learning into PD–LGD–EAD Modeling

Samuel Nyarangi

Independent working paper

August 2026

Abstract

Fintech lenders and digital-asset platforms now originate credit with high-frequency, non-traditional, and—in the on-chain case—publicly observable features. This working paper develops a unified credit-risk architecture in which machine-learning estimators sit inside, rather than beside, the classical PD–LGD–EAD identity used for expected loss, IFRS 9 expected credit loss, and the Basel internal-ratings-based (IRB) capital function. I first restate the loss identity and the Vasicek–Basel asymptotic single-risk-factor formula so that any predictor, linear or not, maps into a probability of default, a loss given default, and an exposure at default. I then specify gradient-boosted and related estimators for those three parameters, with explicit links to scorecards, survival models, and collateral-driven liquidation. For digital assets the same objects are rewritten in loan-to-value and liquidation-time form, using wallet-level and protocol-level features that have begun to appear in the DeFi risk literature. The last part of the paper treats explainability and model risk as first-class constraints: Shapley attributions, stability of rank-order, and the supervisory tests in SR 11-7 and the EBA discussion of machine learning for IRB models. The contribution is architectural rather than empirical. No new loan tape is estimated. The claim is narrower and, I argue, more useful: machine learning improves credit-risk measurement in fintech and digital-asset markets only when its outputs remain identifiable as PD, LGD, and EAD, and only when those outputs can be challenged in the same language supervisors already use.

Keywords: credit risk; fintech; digital assets; machine learning; probability of default; loss given default; explainable AI; DeFi lending.

JEL: G21, G23, G32, C45, C53, O33.

1 Introduction

Credit risk is the chance that a counterparty fails to meet a contractual obligation and that the resulting loss is material. Banks have measured that chance for decades with a small set of objects: the probability of default (PD), loss given default (LGD), and exposure at default (EAD). Those three quantities appear in expected-loss budgeting, in the IFRS 9 expected-credit-loss (ECL) calculation, and in the Basel IRB risk-weight function (Basel Committee on Banking Supervision 2020, 2023; International Accounting Standards Board 2014). The mathematics is older than fintech. What has changed is the data and the function class used to produce the three objects.

Two market developments make the change urgent. First, fintech and big-tech lenders now score borrowers with digital footprints, platform behaviour, and other non-bureau features. Berg et al. (2020) show that even simple digital traces carry default signal. Gambacorta et al. (2024) compare machine-learning models that use non-traditional data with conventional default models at a Chinese fintech and find that the machine-learning specification is better able to predict losses and defaults, especially through nonlinearities. Jagtiani and Lemieux (2019) document related gains from alternative data and machine learning on a U.S. marketplace platform. Second, digital-asset markets have created a second credit book: crypto-collateralized loans, over-collateralized DeFi pools, and—more recently—attempts at under-collateralized, wallet-reputation credit (Auer et al. 2024; Ghosh et al. 2025; Schär 2021). Galaxy Research reports that outstanding loans on DeFi applications reached \$40.99 billion at the end of the third quarter of 2025, with combined DeFi and CeFi crypto-collateralized borrows of \$65.37 billion (Galaxy Research 2025). Those books do not look like a mortgage tape. They do, however, still generate default, recovery, and exposure.

The practical question is therefore not whether machine learning “beats logistic regression on AUC.” Large benchmarking studies have already shown that modern classifiers, and especially ensembles, improve discrimination on credit-scoring data (Baesens et al. 2003; Khandani et al. 2010; Lessmann et al. 2015). The question is whether those gains can be translated into PD, LGD, and EAD without breaking the identities that capital, provisioning, and pricing still use. The European Banking Authority put that question on the table in its November 2021 discussion paper on machine learning for IRB models (European Banking Authority 2021). U.S. supervisors had already required that any model used for material decisions satisfy the documentation, validation, and challenge standards of SR 11-7 (Board of Governors of the Federal Reserve System and Office of the Comptroller of the Currency 2011). Industry surveys confirm that credit risk is among the earliest production uses of AI and machine learning in financial firms (Financial Stability Board 2017; Institute of International Finance and EY 2024).

This paper makes three contributions, all architectural.

First, it writes the mapping from a machine-learning score to PD, LGD, and EAD in a single notation, including the IFRS 9 multi-period discount and the Basel IRB capital function. The point is to stop treating “the model” as a black-box ranker and to treat it as an estimator of three loss components.

Second, it extends the same mapping to digital-asset credit, where the dominant risk is often a collateral liquidation rather than a classical payment default. The loan-to-value process, a liquidation indicator, and a wallet-level on-chain score sit in the same PD–LGD–EAD slots.

Third, it treats explainability and fairness as constraints on admissible estimators, not as optional reports. Fuster et al. (2022) show that machine learning in U.S. mortgage pricing can widen rate disparities. Shapley-based attributions (Lundberg and Lee 2017; Shapley 1953) and standard discrimination and calibration metrics then become part of the model itself, because a PD that cannot be challenged is not a usable IRB or ECL input.

The paper is a working paper, not an estimation paper. I do not introduce a proprietary loan tape or claim a new out-of-sample AUC. Where I write a formula, it is either a standard identity, a named estimator from the literature, or a transparent rewriting of those objects for on-chain collateral. Claim strength is limited to that.

The rest of the paper is organized as follows. Section 2 sets the institutional background. Section 3 states the loss identities. Section 4 specifies machine-learning estimators for the three risk components.

Section 5 extends the architecture to digital assets. Section 6 covers explainability, fairness, and governance. Section 7 discusses validation. Section 8 discusses implications and limits. Section 9 concludes.

2 Institutional setting

2.1 What “credit risk” still means

A lender cares about two events: that the borrower fails, and that the failure costs money. Classical consumer and corporate scoring concentrated on the first event, usually with a binary default flag over a fixed horizon (Altman 1968; Thomas 2000). Basel and IFRS 9 force the second event into the open. Expected loss on a position is

$$EL = PD \times LGD \times EAD, \quad (1)$$

where PD is a probability on a defined horizon, LGD is a loss rate conditional on default, and EAD is the exposure drawn at default (Basel Committee on Banking Supervision 2020). Nothing in (1) requires the three factors to be logistic. Everything in (1) requires them to be defined, comparable across obligors, and stable enough to be challenged.

IFRS 9 replaces the one-period product with a discounted path of expected credit loss (International Accounting Standards Board 2014):

$$ECL = \sum_{t=1}^T PD_t LGD_t EAD_t DF_t, \quad (2)$$

where DF_t is a discount factor and the PD_t are either 12-month or lifetime probabilities according to stage. A machine-learning model that outputs only a rank is not an ECL model. A machine-learning model that outputs a calibrated term structure of PD_t , plus LGD_t and EAD_t , can be.

2.2 Fintech credit

Fintech credit is ordinary credit originated through a digital channel, often with thin or no bureau history. The empirical literature has two consistent messages. Alternative data carry signal: Berg et al. (2020) show that digital-footprint variables predict default even after conventional controls. Machine-learning function classes extract more of that signal than linear scorecards, particularly in stress and in nonlinear regions of the feature space (Bazarbash 2019; Frost et al. 2019; Gambacorta et al. 2024). Those gains are not free. They change who is priced in and who is priced out (Fuster et al. 2022), and they make documentation, challenge, and adverse-action language harder (Board of Governors of the Federal Reserve System and Office of the Comptroller of the Currency 2011; European Banking Authority 2021).

2.3 Digital-asset credit

Digital-asset credit is not one product. Over-collateralized DeFi lending—Aave-style pools and collateralized debt positions—replaces underwriting with a liquidation rule on observable collateral (Auer et al. 2024; Schär 2021). CeFi crypto lenders reintroduce counterparty and operational risk. A third, still thin, layer tries to underwrite wallets rather than people, using on-chain histories as a credit file (Ghosh et al. 2025). Adamyk et al. (2025) survey the risk-management stack around DeFi platforms and note

the growing use of machine learning for anomaly detection and protocol monitoring. The architectural claim of this paper is that all three products still require PD, LGD, and EAD. The definitions change; the identity does not.

3 Loss identities and the capital map

3.1 Default, recovery, and exposure

Fix a horizon $h > 0$ and an obligor i with feature vector $x_i \in \mathbb{R}^p$. Let $D_i(h) = \mathbf{1}\{\tau_i \leq h\}$ be the default indicator, where τ_i is the default time. Then

$$\text{PD}_i(h) \equiv \mathbb{P}(D_i(h) = 1 \mid x_i). \quad (3)$$

Let E_i be contractual exposure and L_i the economic loss if default occurs. Then

$$\text{LGD}_i \equiv \mathbb{E}\left[\frac{L_i}{E_i^*} \mid D_i(h) = 1, x_i\right], \quad \text{EAD}_i \equiv \mathbb{E}[E_i^* \mid x_i], \quad (4)$$

where E_i^* is exposure at the default date (drawn balance plus a credit-conversion term for undrawn lines). Unconditional expected loss is again (1).

A score $s(x)$ is not a PD. A monotone map

$$\text{PD}(h; x) = G_h(s(x)) \quad (5)$$

is a PD only if G_h is a calibrated, horizon-specific function—platt scaling, isotonic regression, or a dedicated probability model. This distinction is elementary and routinely ignored in vendor “AI score” decks.

3.2 Structural default as a special case

The Merton firm-value model (Merton 1974) remains the cleanest structural special case. If asset value V_t follows a geometric Brownian motion and default occurs when $V_T < K$, then

$$\text{PD} = \Phi\left(-\frac{\ln(V_0/K) + (\mu - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right), \quad (6)$$

where Φ is the standard normal distribution function. Machine learning does not replace (6); it replaces the assumption that a low-dimensional diffusion, observed without error, is a sufficient statistic. In fintech the sufficient statistic is a high-dimensional x . In digital-asset lending the analogous state is often a collateral price and an on-chain liquidity path, not a corporate asset value.

3.3 The Basel IRB map

Under the asymptotic single-risk-factor model of Vasicek (2002), which underlies the Basel IRB corporate risk-weight function (Basel Committee on Banking Supervision 2023), unexpected loss for a well-diversified portfolio depends on a Gaussian copula of defaults. For a given PD, LGD, maturity M ,

and asset correlation $\rho(\text{PD})$, the capital requirement per unit of exposure takes the form

$$K = \text{LGD} \left[\Phi \left(\frac{\Phi^{-1}(\text{PD}) + \sqrt{\rho} \Phi^{-1}(0.999)}{\sqrt{1 - \rho}} \right) - \text{PD} \right] \frac{1 + (M - 2.5)b}{1 - 1.5b}, \quad (7)$$

with $b = b(\text{PD})$ the supervisory maturity adjustment and $\rho = \rho(\text{PD})$ the prescribed correlation. Risk-weighted assets are $\text{RWA} = 12.5 K \text{ EAD}$.

Two observations follow immediately. First, any machine-learning PD that enters (7) is a regulatory input, not a marketing score: it must be a long-run, default-weighted probability on the supervisory definition, not an uncalibrated ranking. Second, an improvement in AUC that is achieved by shifting mass in the far left tail can change K by more than it changes expected loss, because (7) evaluates a 99.9% Gaussian quantile. That is why the EBA discussion of machine learning for IRB models concentrates on explainability, representativeness of data, and the economic meaning of the output, not on leaderboard metrics (European Banking Authority 2021).

3.4 A lemma on admissible ML

Proposition 1 (Admissible plug-in). *Let $\widehat{\text{PD}}$, $\widehat{\text{LGD}}$, and $\widehat{\text{EAD}}$ be estimators such that, conditionally on x ,*

$$\mathbb{E}[\widehat{\text{PD}} \widehat{\text{LGD}} \widehat{\text{EAD}} \mid x] = \text{EL}(x).$$

Then the triple is an unbiased plug-in for (1). If any factor is a rank-preserving transform that is not calibrated to the relevant conditional expectation, the product is not an unbiased estimator of expected loss, even if the ranker maximises AUC.

The proof is immediate from the law of iterated expectations and the definition of EL. The proposition is the paper’s organising constraint. Machine learning is allowed to estimate each factor; it is not allowed to replace the product with a nameless score.

4 Machine-learning estimators for the three components

4.1 Probability of default

The workhorse parametric model remains the logistic scorecard

$$\text{logit PD}(x) = \beta_0 + \beta^\top x, \quad \text{PD}(x) = \frac{1}{1 + e^{-(\beta_0 + \beta^\top x)}}. \quad (8)$$

It is still the right baseline. It is not, on current evidence, the best discriminator. Lessmann et al. (2015) update the earlier benchmark of Baesens et al. (2003) and find that heterogeneous ensembles improve on individual classifiers across several real credit-scoring data sets. Khandani et al. (2010) show that nonparametric machine-learning algorithms improve consumer default forecasts relative to standard credit-risk models. Gradient boosting (Friedman 2001) is the function class that now dominates production fintech scoring.

Write the boosting model as an additive expansion in a function class $\mathcal{H}$ of shallow trees,

$$F_M(x) = F_0 + \sum_{m=1}^M \nu h_m(x), \quad h_m \in \mathcal{H}, \quad (9)$$

with shrinkage $\nu \in (0, 1]$. At step m one fits h_m to the negative gradient of a proper scoring rule. For default classification the Bernoulli deviance is

$$\ell(y, F(x)) = -yF(x) + \log(1 + e^{F(x)}), \quad (10)$$

so the pseudo-residual is $y - \text{PD}_F(x)$. Random forests (Breiman 2001) average de-correlated trees instead of adding them sequentially. Either object produces a score $s(x) = F_M(x)$. Calibration (5) then converts the score into $\text{PD}(h; x)$.

Two fintech-specific points matter. First, non-traditional features change the information set, not only the function class. Gambacorta et al. (2024) isolate both effects and find that machine learning plus non-traditional data outperforms traditional models in predicting losses and defaults. Second, discrimination is not the same as economic value. Hand (2009) warns that area under the ROC curve can mislead when class imbalance and decision thresholds matter. A credit model should be reported with a threshold-aware metric—expected loss, or a profit-weighted loss—alongside AUC and the Kolmogorov–Smirnov statistic

$$\text{KS} = \sup_u |F_{s|Y=1}(u) - F_{s|Y=0}(u)|. \quad (11)$$

4.2 Loss given default

LGD is a rate on $[0, 1]$, often with a point mass at full recovery and another near total loss. A linear regression on workout recoveries ignores both. Two machine-learning strategies are standard.

The first is a two-part model. A classifier estimates $\mathbb{P}(\text{LGD} = 0 \mid x)$ or $\mathbb{P}(\text{LGD} = 1 \mid x)$; a bounded regression, beta regression, or a second learner estimates the interior mean. The second is a direct ensemble on the observed loss rate, with a link that respects the unit interval,

$$\widehat{\text{LGD}}(x) = \sigma(F_M^{\text{LGD}}(x)), \quad \sigma(z) = (1 + e^{-z})^{-1}. \quad (12)$$

Downturn LGD, required for IRB, is not the same object as the sample mean. A conservative map

$$\text{LGD}^{\text{DT}}(x) = \widehat{\text{LGD}}(x) + \kappa \widehat{\sigma}_{\text{LGD}}(x) \quad (13)$$

makes the downturn add-on explicit. The constant κ is a policy choice, not a fitting result, and should be documented as such (Basel Committee on Banking Supervision 2020; European Banking Authority 2021).

For digital-asset loans the interior of LGD is often close to a liquidation shortfall, which Section 5 treats directly.

4.3 Exposure at default

For term loans, EAD is close to current drawn balance. For revolving lines,

$$\text{EAD} = B + \text{CCF} \times U, \quad (14)$$

where B is drawn balance, U is undrawn commitment, and CCF is a credit-conversion factor. Machine learning estimates $\text{CCF}(x)$ rather than a segment average, using the same boosting or forest class as for PD, subject to the box constraint $\text{CCF} \in [0, 1]$. In crypto-collateralized lending the analogue of U is unused borrowing power against posted collateral, which can be drawn in minutes. A static CCF is then a poor description; a short-horizon conditional expectation is better.

4.4 Survival and lifetime PD

IFRS 9 lifetime ECL needs more than a 12-month PD. A discrete hazard $h_t(x) = \mathbb{P}(\tau = t \mid \tau \geq t, x)$ yields

$$\text{PD}(0, T; x) = 1 - \prod_{t=1}^T (1 - h_t(x)). \quad (15)$$

Each h_t can be a logistic model or a boosting model on a person-period file. Macro covariates can enter h_t if the lender needs a point-in-time path rather than a through-the-cycle PD. The machine-learning step is again local: it estimates h_t . The identity (15) remains the lifetime map.

5 Digital assets: the same identity, a different state

5.1 Collateral as the state variable

Consider a loan with stablecoin (or fiat) face value D_t , collateral quantity C_t , and collateral mark P_t . The loan-to-value ratio is

$$\text{LTV}_t = \frac{D_t}{C_t P_t}. \quad (16)$$

A protocol or desk liquidates when LTV_t crosses a threshold $\theta \in (0, 1]$, i.e. when

$$P_t \leq P_t^* \equiv \frac{D_t}{C_t \theta}. \quad (17)$$

If the only credit event is a failed liquidation—the collateral cannot be sold at P^* in time—then

$$\text{PD}(\Delta) = \mathbb{P}(P_{t+\Delta} \leq P_t^* \text{ and liquidation fails}). \quad (18)$$

A geometric-Brownian special case,

$$dP_t = \mu P_t dt + \sigma P_t dW_t, \quad (19)$$

gives a closed-form hitting probability analogous to (6). That formula is a benchmark, not a description of crypto prices. Production models replace (19) with a machine-learning or high-frequency volatility model whose state includes realised volatility, order-book depth, stablecoin depeg indicators, and protocol utilisation (Adamyk et al. 2025; Schär 2021).

Conditional on a liquidation, LGD is the shortfall

$$\text{LGD} = \mathbb{E} \left[\max \left\{ 0, 1 - \frac{(1 - \lambda) C_t P^{\text{fill}}}{D_t} \right\} \middle| \text{liquidation} \right], \quad (20)$$

where λ is a fee-and-slippage rate and P^{fill} is the average fill price. EAD is D_t plus any simultaneously drawable amount against residual collateral. Equations (18)–(20) are (1) with a market-risk state. They are not a reason to abandon PD–LGD–EAD.

5.2 Wallet-level underwriting

Under-collateralized digital-asset credit needs a borrower-level PD. Ghosh et al. (2025) propose an on-chain credit risk (OCCR) score: a probabilistic assessment of wallet-level credit risk constructed from observable chain activity. Architecturally the score is another $s(x)$ in (5), with x now containing repayment and liquidation history, leverage cycles, protocol diversity, and related-address clustering. Two caveats belong in the model card, not the appendix. Wallet features are public and therefore gameable. Address clustering is a statistical claim, not an identity. A PD built on a mis-clustered graph is a mis-specified PD.

5.3 A joint conventional–on-chain predictor

Many desks now see both books: a fiat or stablecoin receivable and a crypto collateral sleeve. A stacked predictor

$$s(x) = \alpha s^{\text{off}}(x^{\text{off}}) + (1 - \alpha) s^{\text{on}}(x^{\text{on}}) + g(x^{\text{off}}, x^{\text{on}}) \quad (21)$$

lets the data choose the blend α and a cross-term g . The output still passes through G_h in (5). The stack is how fintech and digital-asset data meet without inventing a fourth risk parameter.

6 Explainability, fairness, and governance

6.1 Why a better AUC can still be an inadmissible model

SR 11-7 requires conceptual soundness, process verification, ongoing monitoring, and effective challenge (Board of Governors of the Federal Reserve System and Office of the Comptroller of the Currency 2011). The EBA discussion paper asks, in IRB language, how a machine-learning algorithm can remain interpretable enough that the risk components are understood, the definition of default is respected, and overrides remain possible (European Banking Authority 2021). A model that cannot answer “why was this obligor a 4.2% PD?” fails those tests even if it wins a Kaggle-style bake-off.

6.2 Shapley attributions as a linear explanation of a nonlinear PD

Let $v(S)$ be the model output—or the calibrated PD—when only features in $S \subseteq N = \{1, \dots, p\}$ are used. The Shapley value of feature j is

$$\phi_j = \sum_{S \subseteq N \setminus \{j\}} \frac{|S|! (|N| - |S| - 1)!}{|N|!} [v(S \cup \{j\}) - v(S)], \quad (22)$$

with $\sum_j \phi_j = v(N) - v(\emptyset)$ (Shapley 1953). Lundberg and Lee (2017) give a unified, computationally tractable family of estimators (SHAP) for this decomposition. For a logistic or boosting PD, one reports ϕ_j in probability units after the calibration map G_h , not only in raw-score units. That is the difference between an explanation of the ranker and an explanation of the risk component.

SHAP does not make a model causal. It makes a prediction additive in a well-defined cooperative-game sense. That is enough for challenge, not enough for a structural claim about why a borrower defaults.

6.3 Fairness as a constraint on $G_h \circ s$

Fuster et al. (2022) study U.S. mortgage pricing and find that machine learning would offer lower rates to groups that already received advantageous rates under a traditional model, while Black and Hispanic borrowers are disproportionately less likely to gain. The result is not a reason to forbid machine learning. It is a reason to treat group-conditional calibration as a design constraint. Write A for a protected attribute used only in monitoring, not as a feature. A minimum requirement is

$$\mathbb{E}[Y \mid \widehat{\text{PD}}, A = a] \approx \widehat{\text{PD}} \quad \text{for each } a, \quad (23)$$

together with a documented rule for what happens when (23) fails. Fintech digital-footprint models are especially exposed, because the footprint can proxy for A even when A is excluded (Berg et al. 2020; Fuster et al. 2022).

6.4 Model risk as an add-on, not a paragraph

Let $\mathcal{M}$ be a set of challenger specifications—logistic, gradient boosting, random forest, and a stacked blend. A simple model-risk add-on on expected loss is

$$\text{EL}^{\text{MR}} = \text{EL}(\widehat{m}) + \eta \max_{m \in \mathcal{M}} |\text{EL}(m) - \text{EL}(\widehat{m})|, \quad (24)$$

with $\eta \in [0, 1]$ a board-level parameter. Equation (24) is not in Basel. It is a way to make disagreement among admissible estimators visible in the same units as (1).

7 Validation

Validation has to speak the language of (1), not only the language of classification.

Discrimination is reported with AUC and KS, with the Hand warning attached (Hand 2009; Lessmann et al. 2015). Calibration is reported with a reliability diagram, the Brier score

$$\text{BS} = \frac{1}{n} \sum_{i=1}^n (\widehat{\text{PD}}_i - Y_i)^2 \quad (25)$$

(Brier 1950), and a Hosmer–Lemeshow-type grouping test (Hosmer and Lemeshow 1980). Stability is reported by vintage and by a population-stability index on the feature distribution. For IRB use, through-the-cycle behaviour and downturn LGD are tested against the supervisory definitions, not against the last twelve months of originations (Basel Committee on Banking Supervision 2020; European Banking

Authority 2021). For IFRS 9, stage assignment and lifetime PD paths are back-tested on accounts that actually rolled (International Accounting Standards Board 2014).

Digital-asset books need two extra tests. First, a liquidation back-test: did $PD(\Delta)$ in (18) match realised liquidations, including failed liquidations? Second, an oracle-and-liquidity test: did P_t and depth at the time of liquidation match the marks used in training? A model that is well calibrated on a quiet tape and silent in a depeg is not validated.

None of these tests require a new theory. They require that the machine-learning object be named as PD, LGD, or EAD before the test is run.

8 Discussion

The literature already supports four statements at citation level. Machine-learning classifiers improve discrimination on credit data relative to older individual algorithms (Khandani et al. 2010; Lessmann et al. 2015). Non-traditional data add default signal in fintech settings (Berg et al. 2020; Gambacorta et al. 2024; Jagtiani and Lemieux 2019). Those gains can reallocate credit across groups (Fuster et al. 2022). Digital-asset credit is large enough, and structurally different enough, that on-chain states and wallet scores now belong in the same conversation as bureau scores (Adamyk et al. 2025; Galaxy Research 2025; Ghosh et al. 2025; Schär 2021). Supervisors have noticed the first three points (Board of Governors of the Federal Reserve System and Office of the Comptroller of the Currency 2011; European Banking Authority 2021; Financial Stability Board 2017; Institute of International Finance and EY 2024).

What the literature does not yet support is a claim I have deliberately not made: that a single production architecture has been shown, on a public multi-lender tape, to dominate IRB and IFRS 9 baselines for both fintech unsecured credit and digital-asset collateralized credit. That paper would need data this paper does not have.

The useful next empirical step is therefore not another AUC table on a single marketplace. It is a three-column comparison—logistic, gradient boosting, stacked off-chain/on-chain—reported in the units of (1), (2), and (7), with SHAP and group-conditional calibration beside the columns. Until that tape exists, the contribution that can be made cleanly is the one made here: keep the identities, change the estimators.

9 Conclusion

Fintech and digital-asset markets changed the feature space of credit. They did not repeal the loss identity. This paper has written machine-learning estimators for PD, LGD, and EAD so that a gradient-boosted score, a wallet reputation model, and a liquidation engine all exit into the same three numbers that expected loss, IFRS 9, and Basel IRB already understand. The mathematics is standard. The discipline is not. A ranker that cannot be calibrated, explained, or challenged is not a credit-risk model, however modern the algorithm.

Two implications follow for practice. Lenders should specify the calibration map G_h and the downturn map for LGD with the same care they specify the boosting depth. Supervisors should evaluate machine-learning credit systems as estimators of (1), not as software products. The research implication is equally plain. The next useful paper is an estimation paper that reports EL, ECL, and K , not only AUC.

References

- Adamyk, B. et al. (2025). “Risk Management in DeFi: Analyses of the Innovative Tools and Platforms for Tracking DeFi Transactions”. In: *Journal of Risk and Financial Management* 18.1, p. 38. DOI: [10.3390/jrfm18010038](https://doi.org/10.3390/jrfm18010038).
- Altman, E. I. (1968). “Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy”. In: *Journal of Finance* 23.4, pp. 589–609.
- Auer, R. et al. (2024). “The Technology of Decentralized Finance (DeFi)”. In: *Digital Finance* 6, pp. 55–95. DOI: [10.1007/s42521-023-00088-8](https://doi.org/10.1007/s42521-023-00088-8).
- Baesens, B. et al. (2003). “Benchmarking State-of-the-Art Classification Algorithms for Credit Scoring”. In: *Journal of the Operational Research Society* 54.6, pp. 627–635.
- Basel Committee on Banking Supervision (2020). *CRE32: IRB Approach: Risk Components*. Tech. rep. Bank for International Settlements.
- (2023). *CRE31: IRB Approach: Risk Weight Functions*. Tech. rep. Bank for International Settlements.
- Bazarbash, M. (2019). *FinTech in Financial Inclusion: Machine Learning Applications in Assessing Credit Risk*. IMF Working Paper WP/19/109. International Monetary Fund.
- Berg, T. et al. (2020). “On the Rise of FinTechs: Credit Scoring Using Digital Footprints”. In: *Review of Financial Studies* 33.7, pp. 2845–2897. DOI: [10.1093/rfs/hhz099](https://doi.org/10.1093/rfs/hhz099).
- Board of Governors of the Federal Reserve System and Office of the Comptroller of the Currency (2011). *Supervisory Guidance on Model Risk Management*. Tech. rep. SR 11-7 / OCC 2011-12. Federal Reserve / OCC.
- Breiman, L. (2001). “Random Forests”. In: *Machine Learning* 45.1, pp. 5–32. DOI: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324).
- Brier, G. W. (1950). “Verification of Forecasts Expressed in Terms of Probability”. In: *Monthly Weather Review* 78.1, pp. 1–3.
- European Banking Authority (Nov. 2021). *Discussion Paper on Machine Learning for IRB Models*. Tech. rep. EBA/DP/2021/04. European Banking Authority.
- Financial Stability Board (2017). *Artificial Intelligence and Machine Learning in Financial Services*. Tech. rep. Financial Stability Board.
- Friedman, J. H. (2001). “Greedy Function Approximation: A Gradient Boosting Machine”. In: *Annals of Statistics* 29.5, pp. 1189–1232. DOI: [10.1214/aos/1013203451](https://doi.org/10.1214/aos/1013203451).
- Frost, J. et al. (2019). *BigTech and the Changing Structure of Financial Intermediation*. BIS Working Paper 779. Bank for International Settlements.
- Fuster, A. et al. (2022). “Predictably Unequal? The Effects of Machine Learning on Credit Markets”. In: *Journal of Finance* 77.1, pp. 5–47. DOI: [10.1111/jofi.13090](https://doi.org/10.1111/jofi.13090).
- Galaxy Research (2025). *The State of Crypto Leverage – Q3 2025*. Tech. rep. Galaxy.
- Gambacorta, L. et al. (2024). “How Do Machine Learning and Non-Traditional Data Affect Credit Scoring? New Evidence from a Chinese Fintech Firm”. In: *Journal of Financial Stability* 73, p. 101284. DOI: [10.1016/j.jfs.2024.101284](https://doi.org/10.1016/j.jfs.2024.101284).
- Ghosh, R. et al. (2025). “On-Chain Credit Risk Score in Decentralized Finance”. In: *arXiv preprint arXiv:2412.00710*. Revised March 2025.
- Hand, D. J. (2009). “Measuring Classifier Performance: A Coherent Alternative to the Area under the ROC Curve”. In: *Machine Learning* 77.1, pp. 103–123. DOI: [10.1007/s10994-009-5119-5](https://doi.org/10.1007/s10994-009-5119-5).

- Hosmer, D. W. and S. Lemeshow (1980). “Goodness of Fit Tests for the Multiple Logistic Regression Model”. In: *Communications in Statistics — Theory and Methods* 9.10, pp. 1043–1069.
- Institute of International Finance and EY (2024). *Annual Survey Report on AI/ML Use in Financial Services*. Tech. rep. IIF / EY.
- International Accounting Standards Board (2014). *IFRS 9 Financial Instruments*. Tech. rep. IFRS Foundation.
- Jagtiani, J. and C. Lemieux (2019). “The Roles of Alternative Data and Machine Learning in Fintech Lending: Evidence from the LendingClub Consumer Platform”. In: *Financial Management* 48.4, pp. 1009–1029. DOI: [10.1111/fima.12295](https://doi.org/10.1111/fima.12295).
- Khandani, A. E., A. J. Kim, and A. W. Lo (2010). “Consumer Credit-Risk Models via Machine-Learning Algorithms”. In: *Journal of Banking & Finance* 34.11, pp. 2767–2787. DOI: [10.1016/j.jbankfin.2010.06.001](https://doi.org/10.1016/j.jbankfin.2010.06.001).
- Lessmann, S. et al. (2015). “Benchmarking State-of-the-Art Classification Algorithms for Credit Scoring: An Update of Research”. In: *European Journal of Operational Research* 247.1, pp. 124–136. DOI: [10.1016/j.ejor.2015.05.030](https://doi.org/10.1016/j.ejor.2015.05.030).
- Lundberg, S. M. and S.-I. Lee (2017). “A Unified Approach to Interpreting Model Predictions”. In: *Advances in Neural Information Processing Systems*. Vol. 30.
- Merton, R. C. (1974). “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates”. In: *Journal of Finance* 29.2, pp. 449–470. DOI: [10.1111/j.1540-6261.1974.tb03058.x](https://doi.org/10.1111/j.1540-6261.1974.tb03058.x).
- Schär, F. (2021). “Decentralized Finance: On Blockchain- and Smart Contract-Based Financial Markets”. In: *Federal Reserve Bank of St. Louis Review* 103.2, pp. 153–174. DOI: [10.20955/r.103.153-74](https://doi.org/10.20955/r.103.153-74).
- Shapley, L. S. (1953). “A Value for n -Person Games”. In: *Contributions to the Theory of Games* 2, pp. 307–317.
- Thomas, L. C. (2000). “A Survey of Credit and Behavioural Scoring: Forecasting Financial Risk of Lending to Consumers”. In: *International Journal of Forecasting* 16.2, pp. 149–172.
- Vasicek, O. A. (2002). “The Distribution of Loan Portfolio Value”. In: *Risk* 15.12, pp. 160–162.