Computer and Decision Making

An International Journal

www.comdem.org
eISSN: 3008-1416



IPO Awareness and Retail Investors' Subscription Preferences: An Explainable Ordinal Machine Learning Framework

Utkarsh Tikku[1], Dr. Sarita Agrawal[2], Dr. Happy Agrawal[3]

[1]Research Scholar, Department of Management Studies, Vivekananda Global University, Jaipur, Email ID: utkarshtikku.ibs@gmail.com, ORCID ID: https://orcid.org/0009-0006-2802-45672
[2]Associate Professor, Department of Management Studies, Vivekananda Global University, Jaipur, Rajasthan, Email ID: saritaagr.research@gmail.com, ORCID ID: https://orcid.org/0009-0001-7206-9387
[3]Assistant Professor, Department of Business Administration, St. Xavier's College Jaipur, Rajasthan. Email ID: guptahappy29@gmail.com, ORCID ID: https://orcid.org/0000-0001-7247-668X

ARTICLE INFO

Article history:
Received 19 July 2026
Received in revised form 08 September 2026
Accepted 10 October 2026
Available online 29 October 2026

Keywords: IPO Awareness; Retail Investors; Ordinal Machine Learning; Explainable Artificial Intelligence; Financial Decision Support.

ABSTRACT

Initial public offering (IPO) subscription preferences of retail investors must be predicted using computational methods that take investor heterogeneity, ordinal outcomes, and an informational characteristic into account. This study constructed an explainable ordinal machine-learning model to compare the incremental predictive power of IPO awareness and to test the reliability of financial decision-support. A total of 176 UAE retail investors and 1,936 respondent-IPO observations pertaining to 11 IPOs were analysed. Respondent-grouped nested cross-validation was used for the evaluation of Awareness-Enhanced Ordinal XGBoost (AE-OXGB v4) and Awareness-Enhanced Ordinal CatBoost (AE-OCB). Predictive accuracy, interpretability, generalizability and probability calibration were assessed by awareness ablation, subgroup analysis, leave-one-IPO-out validation, permutation feature importance and temperature scaling. AE-OCB had lower ordinal mean absolute error (MAE = 0.9029) compared to AE-OXGB v4 (0.9520) but overall discrimination of the classifications was low. When we only considered features without awareness, MAE for CatBoost was 0.8838 and for XGBoost was 0.9070, which shows that there was not enough incremental predictive power. Risk attitude and focal IPO awareness were two small, unstable positive permutation importance. The transfer performance was found to be heterogeneous from the results of Cross-IPO evaluation. Logarithmic loss decreased from 1.5508 to 1.4068 and the probability oriented performance improved, with the macro-F1 decreasing, as a result of the temperature scaling. The results suggest that informational relevance does not necessarily go hand-in-hand with improved prediction. The proposed framework integrates the potential of the ordinal modelling, the explainability and reliability assessment and provides methodological insights for financial decision support. But the amount of predictive discrimination (and sample size) will need to be validated independently for operational deployment.

1. Introduction

Initial public offerings (IPOs) play an important role in the formation of capital, development of financial markets, and in retail investment. The information accessibility, financial knowledge, perceived risk and behavioural characteristics affect the investors' subscription preferences. Awareness of the IPO is

* *Corresponding author.*
E-mail address: utkarshtikku.ibs@gmail.com

crucial as it serves as an informational factor, meaning that investors cannot assess the investment potential of an IPO until they are aware of it. But being aware does not mean being willing to pay for a subscription: financial decisions are subject to heterogeneous expectations and considerations of risk. Liu et al. [1] studied the relationship between limited investor attention and the underpricing of IPOs, Han and Zhang [2] studied the relationship between social media analysts, investor attention and IPO pricing. Mengoli and Pattitoni [3] then investigated spillover effects to attention with respect to IPO activity. The purpose of these studies is to show the importance of investors' attention, but in no way directly determine the increment of its contribution in predicting subscription preference.

With the improvement of financial analytics, there has been an increasing focus on the use of machine-learning in solving complex investment prediction problems. Pirayesh Neghab et al. [4] studied machine learning models for pricing IPOs, and Warkulat and Pelster [5] and Münster et al. [6] explored the link between information disclosure and retail investor behavior. However, there are separate financial outcomes with IPO pricing, realized trading activity and subscription preferences. Subscription likelihood represents a investor's expressed interest in subscribing, and is typically an ordered response option.

Standard multiclass classification methods fail to consider the preference ordering aspect, hence demonstrating the necessity of accounting for nonlinear behavioral interactions and IPO-specific factors. Heterogeneity of investors presents another difficulty for making predictions since risk attitudes, financial literacy, demographics, and investment experience affect financial decisions. Overconfidence in household financial decisions was analyzed by Xu et al. [7], whereas Lawrence et al. [8] studied the effect of gender on financial literacy and overconfidence. The mentioned works prove that behavioral and informational variables are essential. However, the awareness variables might be redundant and have heterogeneous effects, which should be empirically estimated.

Finally, there is a need to develop explainable and reliable financial decision support systems. Although boosting models (XGBoost and CatBoost) account for interactions effectively, validation is required. Issues related to interpretability were highlighted by Weber et al. [9] and Černevičienė and Kabašinskas [10]. However, the percentage of correct classification cannot prove the reliability of decisions. Inadequate model interpretation due to ordinal prediction errors, unstable feature contributions, lack of generalization, and poorly calibrated probabilities. Therefore, the following should be incorporated into computational frameworks: Ordinal performance evaluation, Feature importance, Validation across contexts, Probability calibration.

Even though there are advances in IPO analytics and explainable financial prediction, there are three methodological gaps that need investigation. First, the existing literature on computational research that focuses on IPOs has tended to focus on pricing and market-level results, while integrated prediction of individual ordinal subscription preferences is relatively underdeveloped. Second, there is limited evidence on the incremental predictive ability of IPO awareness when combined with demographic and behavioral variables. Third, repeated investment observations and variation among IPO offerings call for separate validation strategies to gauge the extent of generalization, while avoiding overstating the extent of predictive reliability. In addition, there is a need for explainability and probability calibration to work in conjunction with classification performance when assessing financial decision-support applications.

To fill those gaps, this study proposes an explainable ordinal machine-learning model based on secondary survey data for 176 retail investors in the United Arab Emirates who participated in 11 IPOs with a total of 1,936 respondent–IPO observations. Cumulative-threshold classification is used for implementing two approaches: Awareness-Enhanced Ordinal XGBoost (AE-OXGB v4), and Awareness-Enhanced Ordinal CatBoost (AE-OCB). Their contribution is in terms of integrated application and evaluation, not the invention of new boosting algorithms. In the case of repeated observations, respondent-grouped nested cross-validation prevents information leakage. Awareness Ablation and Subgroup Analysis tests incremental predictive information and investor heterogeneity. Leave-one-IPO-out validation verifies transferability between offerings, held-out permutation importance and

temperature scaling measure model explainability and probabilistic reliability.
The study has three research aims:

1. To create and contrast awareness enhanced ordinal XGBoost and CatBoost models to predict the likelihood of IPO subscription by retail investors.
2. To assess the contribution of the predictors IPO awareness and investor characteristics using ablation analysis, subgroup assessment and explainable AI.
3. Cross-validation, cross-IPO testing and probability calibration to evaluate model generalizability and the reliability of the model to support decisions.

The study provides an integrated computational framework that separates the informational relevance, incremental predictive utility, and probability reliability of predicting IPO subscription preferences. It assesses the trustworthiness of algorithmic forecasts for use in financial decision-making. The analysis is based on stated preferences, however, and not actual investment behaviour, of UAE retail investors. The observational design does not allow for causal inferences and cross-IPO validation cannot prove generalizability to new investors. These constraints help to frame the predictive role of the framework and its use in investment decision making.

2. Literature Review

Financial literacy, risk tolerance, behavioural biases and information processing capacities influence retail investment decisions. Wu and Zhang [11] studied the trading behaviour of retail investors, and found evidence of market-timing and clustering. Aman et al. [12] explored the financial literacy pertaining to mutual fund investment, and Oehler et al. [13] explored the relationship between financial literacy and risky asset participation. These studies have confirmed that financial knowledge is a relevant determinant of investment behaviour, but make no distinction between knowledge of specific securities or stated subscription intentions.

Sajid et al. [14] explored the relationship between financial literacy, confidence and financial behaviour, showing the significance of the psychological mechanisms. Kuzubaş and Saltoğlu [15] studied the effects of risk attitudes based on surveys and portfolio risk in the context of financial prediction and thus contributed to the inclusion of individual risk preferences. However, the concepts of financial literacy, investor attention and awareness of IPO are isolated concepts. The literacy level is a measure of financial knowledge, attention is a measure of the allocation of information, and the awareness of an IPO is a measure of the recognition of a specific offering. Actual trading behaviour also differs from reported preferences for subscriptions. Therefore, awareness needs to be assessed in combination with investor characteristics to identify its incremental predictability, not that awareness necessarily results in subscription willingness.

In financial data, there may be nonlinear relationships, interactions between predictors, and even heterogeneous predictors, which machine learning approaches can handle. A study by Chen et al. [16] explored interpretable machine learning in the context of imbalanced credit scoring, focusing on the need to assess the prediction performance under non-uniform class distribution. Liu et al. [17] studied the credit assessment with interpretable ensemble learning, which highlighted the importance of combining feature-level explanations with predictive models.

Nallakaruppan et al. [18] created an explainable AI framework for credit evaluation and Klein and Walther [19] explored progress in financial explainability. While these contributions set the stage for transparent financial prediction, their main use is in contexts other than ordinal IPO subscription modelling.

Feature importance also poses challenges in methodology. Predictors may be correlated with each other and therefore may not be as important individually, but act together to provide a prediction. Perturbation-based explanations can be made more complicated by the use of engineered interactions between variables, since the perturbation of one may lead to an inconsistent set of other variables. Therefore, permutation importance should be viewed as the reliance of the model as a whole and not as causal importance or as a measure of independent economic significance under a given perturbation

procedure. To provide a complementary evidence, the effects of ablating awareness-related information from the modelling configuration are examined in relation with the performance changes. When the subscription to the IPO is concerned, there are ordered preference categories which need ordered-preserving predictive methods. Buczak et al. [20] contrasted tree ensembles with parametric methods for ordinal prediction, which is useful for assessing nonlinear learning methods of ordered outcomes. Binotto and Delgado [21] discussed ordinal classification measures, emphasizing the need to take account of the distances between response categories into account.

For this, the mean absolute error (MAE) of the categories is used, which penalizes more heavily larger errors of the classification than errors within adjacent categories. In addition to classification accuracy, quadratic weighted Cohen's kappa evaluates ordinal agreement, taking into account chance agreement and penalizing larger disagreements in categories.

Haas and Hüllermeier [22] explored cumulative-threshold approaches to ordinal classification by binary reduction, where uncertainty was of interest. Delgado [23] studied label dependent losses and focused on the effect of ordering outcomes in predictive evaluations. The methodological justification for ordinal XGBoost and CatBoost implementations is given by these contributions, all in all.

However, the performance of the prediction is dependent on the design of validation. The independence of the observations in ordinary record-level partitioning is violated by repeated observations from individual investors. Respondent-grouped nested cross-validation uses different groups of investors for training and evaluating the model as well as different hyperparameter values for the outer testing. Leave-one-IPO-out validation tackles another question: Do learned predictive relationships carry over across offerings? These are complementary procedures that differentiate between evaluation of unseen-investors in observed offerings and evaluation of transfer of offerings by unseen.

Financial DSS needs discriminative predictions and reliable probability estimates. Silva Filho et al. [24] discussed methods of calibrating classifiers, focusing on how close the predicted probabilities are to the frequency of outcomes. The issue with discrimination is about the separation between the various categories of outcomes, while the issue with calibration is about the reliability of the probabilities given. A model can thus have a better logarithmic loss or Brier score than macro-F1 or balanced accuracy. The probability adjustment can be done using temperature scaling without re-training the prediction models.

Huang [25] explored feature importance via residual permutation testing and provided methodological considerations when assessing predictor relevance. Both conventional held-out permutation importance and residual permutation tests underscore the importance of clear and unambiguous perturbation methods and careful interpretation of feature importance.

A review of the literature identified three inter-related gaps. Very little integrated assessment of IPO awareness as an incremental predictor of ordinal subscription preferences has been done, controlling for investor characteristics. Second, the studies of financial predictions do not always integrate ordinal learning, awareness ablations, investor heterogeneity and cross offering validations. Third, the explainability, ordinal discrimination, and probability calibration of models are often tested independently, and not all three are tested to assess the reliability of the model for decisions.

Table 1. Comparative Synthesis of Related Literature and Research Gaps

Study	Research focus	Methodological approach	Principal contribution	Gap relevant to this study
Wu and Zhang [11]	Retail trading behaviour	Empirical financial analysis	Investor trading patterns	IPO subscription preferences
Aman et al. [12]	Financial literacy	Empirical investment analysis	Literacy-related investment behaviour	IPO-specific awareness
Kuzubaş and Saltoğlu [15]	Risk attitudes	Survey-based analysis	Risk-preference measurement	Ordinal subscription prediction
Chen et al. [16]	Credit scoring	Interpretable ML	Explainability in financial classification	IPO-specific ordinal outcomes
Liu et al. [17]	Credit prediction	Interpretable ensembles	Financial predictive modelling	Awareness ablation
Buczak et al. [20]	Ordinal prediction	Comparative modelling	Evaluation of ordinal learners	IPO decision-support application
Haas and Hüllermeier [22]	Ordinal uncertainty	Binary reduction	Uncertainty assessment	Cross-IPO transfer
Silva Filho et al. [24]	Probability calibration	Methodological review	Calibration assessment	Integrated ordinal reliability
Huang [25]	Feature importance	Residual permutation testing	Predictor relevance assessment	IPO awareness contribution

Accordingly, the present study integrates respondent-grouped ordinal XGBoost and CatBoost, awareness ablation, investor subgroup analysis, cross-IPO generalization, held-out permutation importance, and temperature scaling. Its contribution lies in systematically evaluating whether IPO awareness adds predictive information while distinguishing ordinal accuracy, feature reliance, generalization, and probabilistic reliability within an integrated financial decision-support framework.

3. Research Methodology

3.1. Research Design and Analytical Framework

A quantitative (secondary data driven) ordinal machine learning approach was adopted in this study to predict the initial public offering (IPO) subscription preference of retail investors. Respondent-grouped nested cross-validation (CV) was used to evaluate Awareness-Enhanced Ordinal XGBoost (AE-OXGB v4) and Awareness-Enhanced Ordinal CatBoost (AE-OCB). The analytical workflow included data preprocessing, awareness-feature engineering, hyperparameter optimization, model comparison, awareness ablation, subgroup evaluation, explainability and probability calibration. A leave-one-IPO-out process was used to evaluate cross-offering generalization.

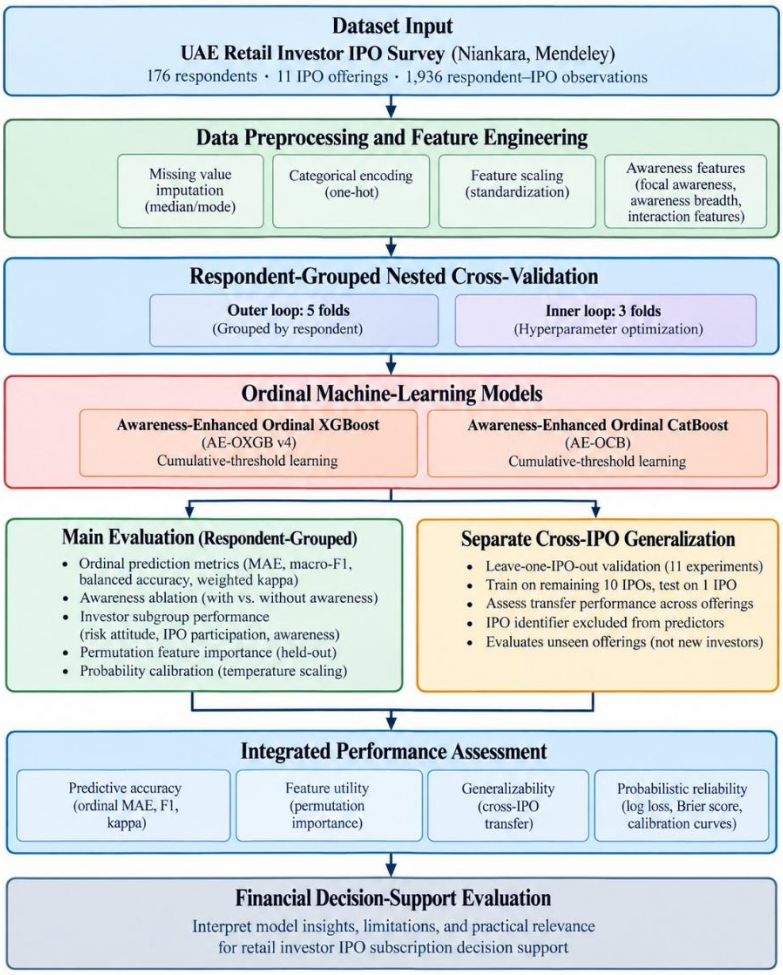


Fig. 1. Proposed Workflow of the Explainable Ordinal Machine-Learning Framework for IPO Subscription Prediction.

3.2. Dataset and Variable Operationalization

The data set of UAE retail investors for this study was obtained from Niankara [26] and consisted of 176 investors who assessed 11 IPOs. Restructuring provided 1,936 respondent–IPO observations, which were repeated observations by the same investors, not independent observations. The respondent identifiers were passed through to grouped validation. Subscription likelihood was measured as four ordered categories ($Y \in \{1,2,3,4\}$) where higher values reflect increasing stated likelihood of subscription. Original IPO awareness responses (QB1: 1 = Yes; 2 = No) were converted to binary indicators (1 = aware; 0 = unaware). Predictors were age, gender, education, occupation, income, investment training, past IPO participation, risk attitude, financial-market knowledge, past investment losses, and IPO identifiers. To reduce leakage of the target, downstream responses related to subscription were excluded.

3.3 Data Preprocessing and Feature Engineering

Numerical predictors underwent median imputation and standardization, whereas categorical variables received mode imputation and one-hot encoding. Only training partitions were used to fit transformations. Four feature configurations were explored: awareness exclusion, focal awareness, awareness breadth, and awareness interactions.

Awareness breadth was defined as:

$$B_{ij} = \sum_{\substack{k=1 \\ k \neq j}}^{11} A_{ik}$$

where A_{ik}represents respondent i's awareness of IPO k. The focal offering was excluded to prevent

duplication. Interaction features combined focal awareness with risk attitude and previous IPO participation, permitting heterogeneous predictive relationships.

3.4 Awareness-Enhanced Ordinal Learning Models

Both approaches adapted established gradient-boosting algorithms through cumulative-threshold decomposition. Three binary classifiers estimated:

$$q_k(x) = P(Y > k \mid x), \quad k = 1,2,3$$

Cumulative probabilities were constrained to satisfy $q_1 \geq q_2 \geq q_3$. Four-category probabilities were reconstructed as:

$$p(x) = [1 - q_1,\ q_1 - q_2,\ q_2 - q_3,\ q_3]$$

$$\hat{y} = \min\left\{k: \sum_{c=1}^{k} p_c\,(x) \geq 0.5\right\}$$

This formulation preserved category ordering while accommodating nonlinear predictor relationships.

3.5 Nested Cross-Validation and Model Evaluation

Five outer and three inner stratified group folds separated respondents during model training, hyperparameter selection, and evaluation. Twenty randomized configurations were assessed per model family and feature configuration.

XGBoost optimization examined tree depth (2–5), learning rate (0.015–0.12), estimators (100–650), minimum child weight (1–10), row and column subsampling (0.70–1.00), gamma (0–1.5), L1 regularization (0–5), and L2 regularization (1–18). CatBoost optimization considered depth (2–5), learning rate (0.02–0.11), iterations (120–500), and L2 leaf regularization (3–30).

Ordinal mean absolute error was the primary selection criterion:

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |\,y_i - \hat{y}_i|$$

Secondary metrics comprised macro-F1, balanced accuracy, quadratic weighted Cohen's kappa, multiclass Brier score, and logarithmic loss. An inner-validation-selected strategy was additionally assessed.

3.6 Awareness Ablation and Uncertainty Assessment

Awareness-enhanced configurations were compared with awareness-excluded counterparts using corresponding selected hyperparameters. Incremental predictive contribution was calculated as:

$$\Delta MAE = MAE_{without} - MAE_{with}$$

Positive differences signified better performance after awareness inclusion. The uncertainty was measured with 600 paired, within-investor, bootstrap resamples, and 95% confidence intervals were calculated using the 2.5th and 97.5th percentiles. The intervals for both XGBoost and CatBoost were zero, meaning that there was no statistically significant incremental improvement in prediction. These comparisons were made for predictive purposes, not causal.

3.7 Investor Heterogeneity and Cross-IPO Generalization

Subgroup performance was assessed for each of the three subgroups: risk attitudes, prior IPO participation, and focal-awareness (with a minimum of 15 respondents per subgroup). Offering-level transfer was a part of 11 leave-one-IPO-out experiments. Ordinal XGBoost was trained on 10 offerings and tested on the offering that was not included in the training set for each experiment. Transfer predictors did not include IPO identifiers. This procedure did not consider independent investor populations but unseen offerings.

3.8 Explainable AI and Feature Importance

Held-out permutation importance assessed model reliance on 13 predictors:

$$\mathrm{PI_j} = \mathrm{MAE}\left(\mathrm{X_{\pi_j}}\right) - \mathrm{MAE(X)}$$

Here, $\mathrm{X_{\pi_j}}$denotes validation observations with predictor jpermuted. Positive values indicated increased prediction error, while negative values indicated reduced error following perturbation. Mean importance and standard deviation summarized fold-wise estimates. Correlated predictors limited independent attribution; importance was not interpreted causally.

3.9 Probability Calibration and Decision-Support Evaluation

Temperature scaling adjusted predicted probabilities:

$$\tilde{\mathrm{p}}_{\mathrm{ik}}(\mathrm{T}) = \frac{\mathrm{p}_{\mathrm{ik}}^{1/\mathrm{T}}}{\sum_{\mathrm{c}=1}^{4} \mathrm{p}_{\mathrm{ic}}^{1/\mathrm{T}}}, \qquad \mathrm{T} > 0$$

The outer-test labels were not used for the optimization of the temperature parameter, which is done in the form of predictions in the internal training partition. Logarithmic loss, Brier score and reliability curves were used to assess calibration. Prediction reliability was evaluated using the IPO-level MAE and categorical correctness. These were declared subscription preferences and not investment deals or investment gains.

4. Results and Analysis

4.1. Respondent Characteristics and IPO Awareness

Overall, the data contained 176 UAE investors and 1,936 observations of the investors' responses to the IPOs of 11 offerings. The majority of respondents had no prior experience of participating in IPOs (78.41%) and were not very knowledgeable about the market (68.75%). Salik had the highest awareness rate (56.82%) while Fertiglobe had the lowest (21.02%). All offerings had a median subscription-likelihood score of 2.

Table 2. Respondent Demographic and Investment Characteristics (N = 176)

Characteristic	Category	n	%
Gender	Male	106	60.23
	Female	70	39.77
Education	Bachelor's	51	28.98
	PhD	45	25.57
	High school	37	21.02
	Master's	26	14.77
	Diploma	17	9.66
Previous IPO	Yes	38	21.59
	No	138	78.41
Risk attitude	Highly averse	45	25.57
	Averse	49	27.84
	Neutral	48	27.27
	Taker	34	19.32
Market knowledge	Limited	121	68.75
	General	41	23.30
	Detailed	14	7.95

Table 3. IPO Awareness and Subscription Likelihood

IPO	Awareness (%)	Mean	Median
ADNOC Drilling	50.00	2.381	2
AD Ports	35.23	2.233	2
Bayanat	27.27	1.972	2
Borouge	33.52	2.040	2
Burjeel	48.86	2.131	2
DEWA	52.27	2.347	2
Empower	24.43	1.949	2
Fertiglobe	21.02	1.847	2
Salik	56.82	2.369	2
Tecom	28.41	1.943	2
YahSat	26.70	1.989	2

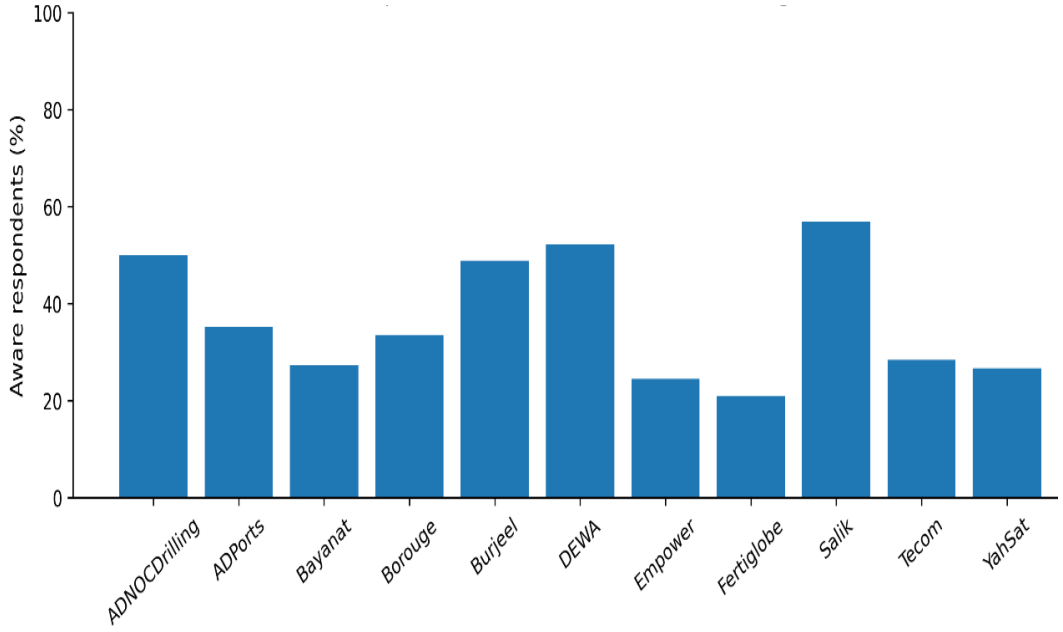


Fig. 2. IPO Awareness Distribution Across Eleven UAE Public Offerings.

4.2. Comparative Ordinal Model Performance

AE-OCB achieved the lowest MAE (0.9029), outperforming AE-OXGB v4 (0.9520) by 5.15%. However, low macro-F1, balanced accuracy, and quadratic weighted kappa indicated limited absolute predictive discrimination (Table 3).

Table 4. Comparative Ordinal Model Performance

Model	F1 ↑	BA ↑	MAE ↓	κ ↑	Brier ↓	Log loss ↓
AE-OXGB v4	0.2062	0.2343	0.9520	0.0521	0.7962	1.9583
AE-OCB	0.2333	0.2767	0.9029	0.0400	0.7599	1.5426
Inner-selected	0.2248	0.2660	0.9158	0.0379	0.7641	1.5508

Note: F1 = macro-F1; BA = balanced accuracy; κ = quadratic weighted kappa.

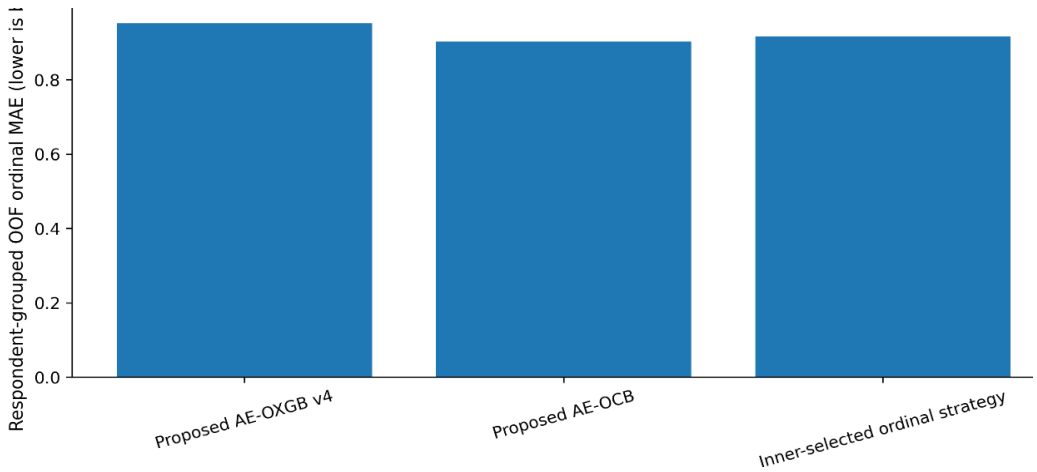


Fig. 3. Comparative Ordinal Prediction Error Across Evaluated Models.

4.3. Awareness Ablation and Uncertainty Assessment

In both model families, there was a reduction of MAE in the one condition in which the students were instructed to exclude awareness: Table 4. Paired respondent-level bootstrap analysis (600 resamples) yielded ΔMAE = −0.0449 (95% CI: −0.0930 to 0.0011) for XGBoost and −0.0191 (95% CI: −0.0661 to 0.0243) for CatBoost. For both, the interval contained 0, which means there are no significant differences.

Table 5. Awareness Ablation Performance

Model	Awareness	F1	MAE ↓	Brier ↓	Log loss ↓
XGBoost	Excluded	0.2157	0.9070	0.7636	1.8023
	Included	0.2062	0.9520	0.7962	1.9583
CatBoost	Excluded	0.2259	0.8838	0.7450	1.4503
	Included	0.2333	0.9029	0.7599	1.5426

Note: ΔMAE = MAE(excluded) − MAE(included); reduced models reused selected hyperparameters.

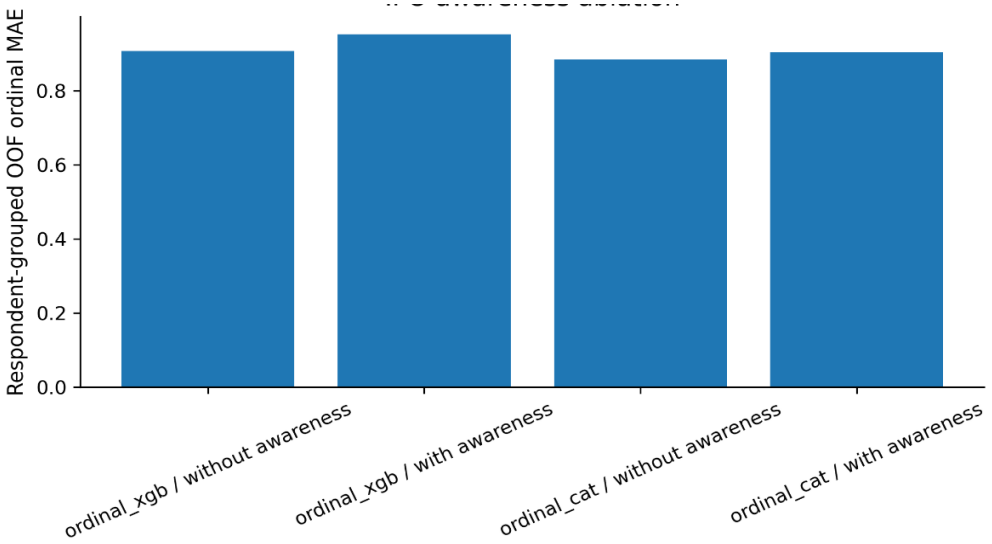


Fig. 4. Ordinal MAE Changes Following Awareness Inclusion.

4.4. Investor Subgroup Performance

Prediction errors were not uniform among different types of investors. The lowest MAE was for the risk-neutral respondents (0.8466), while higher MAE was for IPO-aware respondents (1.0112) as compared to IPO-unaware respondents (0.8603). Differences were descriptive and inferential significance testing were not performed.

Table 6. Investor Subgroup Prediction Performance

Subgroup	Investors	Obs.	F1	MAE ↓
Highly risk-averse	45	495	0.1629	0.9899
Risk-averse	49	539	0.2192	0.8738
Risk-neutral	48	528	0.2183	0.8466
Risk-taker	34	374	0.1861	0.9759
Previous IPO: No	138	1,518	0.2045	0.9269
Previous IPO: Yes	38	418	0.1756	0.8756
IPO unaware	163	1,224	0.2152	0.8603
IPO aware	134	712	0.1661	1.0112

Note: Awareness-group investor counts overlap.

4.5. Cross-IPO Generalization

Heterogeneous transfer performance was shown with leave-one-IPO-out testing. The lowest MAE (0.4261) was recorded by Burjeel whereas the highest was recorded by ADNOC Drilling (0.8239). This assessment considered unrecognized offerings, not independent investor groups.

Table 7. Leave-One-IPO-Out Generalization

Held-out IPO	F1 ↑	BA ↑	MAE ↓	κ ↑
ADNOC Drilling	0.3433	0.4051	0.8239	0.3474
AD Ports	0.6301	0.6250	0.4318	0.6780
Bayanat	0.4146	0.4360	0.5568	0.4083
Borouge	0.2345	0.3255	0.8125	0.0900
Burjeel	0.6286	0.6297	0.4261	0.6854
DEWA	0.3826	0.4273	0.7159	0.3746
Empower	0.2203	0.2969	0.7443	0.1404
Fertiglobe	0.2145	0.2908	0.7443	0.1067
Salik	0.4147	0.4699	0.6932	0.4187
Tecom	0.2250	0.2815	0.6818	0.1333
YahSat	0.2629	0.3267	0.6989	0.1094

4.6. Explainable AI and Feature Importance

The small positive permutation importance of risk attitude (0.0110) and focal awareness (0.0026) were similar in size. The other predictors had negative estimates. Selected were different types of awareness breadth in folds 1–3, and in folds 4–5, different types of focal-awareness configurations. The fold coverage of the importance estimates varies.

Table 8. Held-Out Permutation Feature Importance

Predictor	Mean ΔMAE	SD	Folds
Risk attitude	0.0110	0.0442	5
Focal awareness	0.0026	0.0104	5
Gender	−0.0005	0.0238	5
Age	−0.0014	0.0420	5

IPO identifier	−0.0016	0.0064	5
Investment training	−0.0073	0.0085	5
Income	−0.0116	0.0450	5
Market knowledge	−0.0119	0.0185	5
Education	−0.0122	0.0245	5
Occupation	−0.0130	0.0436	5
Previous IPO	−0.0174	0.0264	5
Previous market loss	−0.0195	0.0337	5
Awareness breadth	−0.0343	0.0442	3

Note: Positive ΔMAE indicates increased error after permutation, not causality.

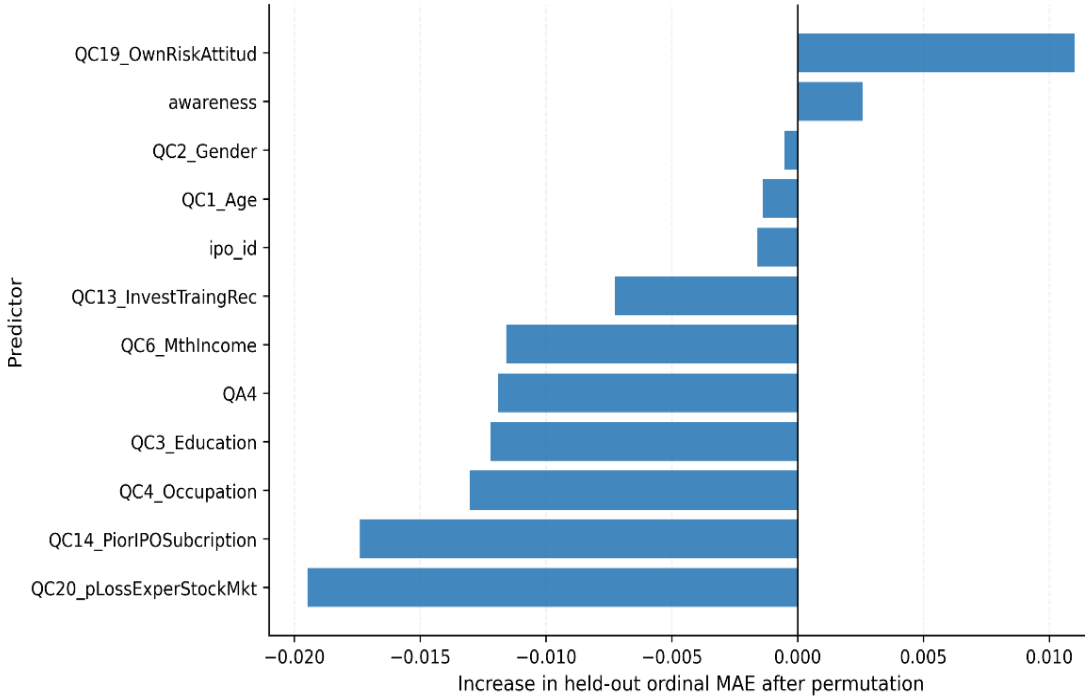


Fig. 5. Held-Out Permutation Feature Importance.

4.7. Probability Calibration and Decision-Support Evaluation

Temperature scaling resulted in a reduction of log loss by 9.29%, and an increase in the Brier score and MAE. Macro-F1 and balanced accuracy, however, decreased, reflecting a calibration–discrimination trade-off.

Table 9. Performance Before and After Calibration

Metric	Before	After
Macro-F1 ↑	0.2248	0.1599
Balanced accuracy ↑	0.2660	0.2532
MAE ↓	0.9158	0.8714
Quadratic κ ↑	0.0379	0.0544
Brier ↓	0.7641	0.7423
Log loss ↓	1.5508	1.4068

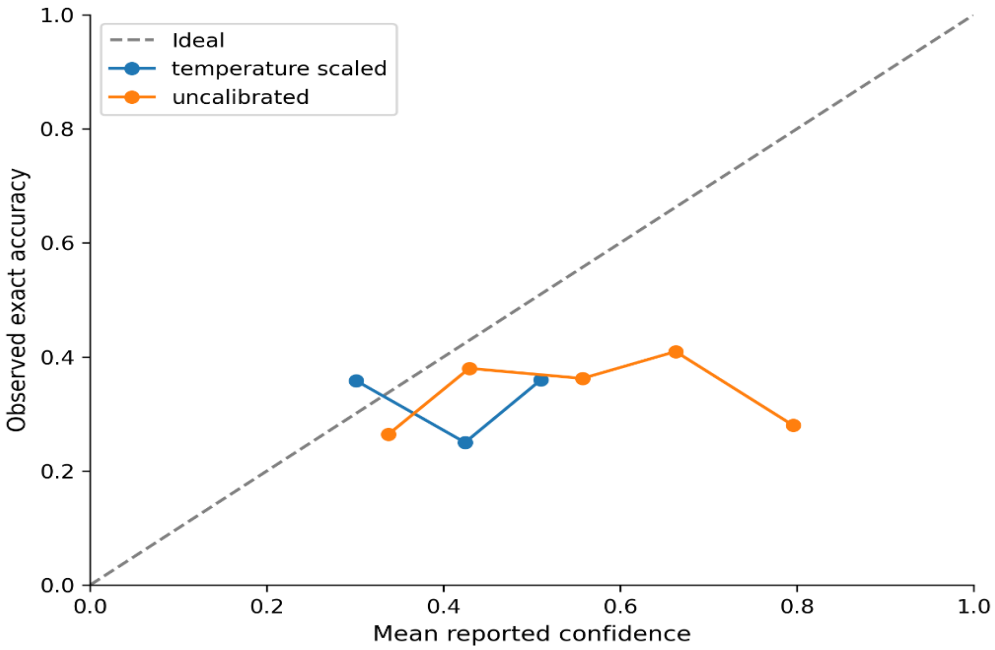


Fig. 6. Reliability Curves Before and After Temperature Scaling.

In grouped-validation IPO-level evaluation, Tecom achieved the highest categorical accuracy (41.48%) and lowest MAE (0.7443); ADNOC Drilling recorded the lowest accuracy (24.43%) and highest MAE (1.1136). These findings are distinct from cross-IPO transfer results. Overall, predictive utility remained limited despite relative model and calibration improvements.

5. Discussion

Results indicate that ordinal machine learning has methodological relevance and prediction limitations regarding retail investors' IPO subscription preferences. Within the configurations tested, AE-OCB resulted in the lowest ordinal MAE value (0.9029), which means that it performed relatively better than AE-OXGB v4 (0.9520). This variation may be due to the differences in categorical feature processing, regularization, and optimization, which were not experimentally separated. However, there is limited discrimination and ordinal agreement beyond chance for AE-OCB based on macro-F1 (0.2333) and quadratic weighted kappa (0.0400) respectively. Therefore, relative performance advantages are NOT to be interpreted as an operational predictive advantage. The results are not only applicable to computational finance applications, but also to individual investor-level subscription preferences, not to pricing outcomes of IPOs [4] [16].

We did not find much incremental predictive value of ablation of awareness. The optimized XGBoost MAE after removing predictors that affect awareness dropped from 0.9520 to 0.9070 and CatBoost MAE after removing dropped to 0.9029 to 0.8838. However, paired respondent-level bootstrap confidence intervals included zero for both XGBoost (ΔMAE = −0.0449; 95% CI: −0.0930 to 0.0011) and CatBoost (ΔMAE = −0.0191; 95% CI: −0.0661 to 0.0243). This observed decrease was not statistically significant, therefore. Although investors' interest may create market behaviour for IPO markets [1]–[3], it is not a requirement for its informational relevance to be useful. These findings may be explained by redundant awareness representations, measurement noise, and by heterogeneous behavioural relationships, but these processes are yet to be verified.

There was heterogeneous predictive reliability, with investor subgroup analysis. The MAE of risk neutral respondents is less than that of the highly risk averse investors (0.8466 vs 0.9899) and the IPO aware observations have higher MAE than the unaware observations. The differences can be explained as differences in distributions of response/complexity and/or behaviour/complexity. Descriptive comparisons do not measure the importance of differences, however, and cannot identify cause and effect. Both risk attitude (0.0110) and focal awareness (0.0026) showed small positive effects with

significant variation between the folds. However, the importance values are generally negative, which means that the predictors are not very stable, particularly features that are correlated [9][25].
The cross-IOPO evaluation results revealed that there is a difference in the transfer performance, with a MAE of 0.4261 for Burjeel, and 0.8239 for ADNOC Drilling. Variability in performance may be due to the characteristics of the offering and the distribution of preferences, which need to be explored further. Leave one IPO out validation is not about generalization to independent investors, but about transferability across offerings. The log loss was reduced from 1.5508 to 1.4068 and the Brier score from 0.7641 to 0.7423, while macro-F1 was lowered from 0.2248 to 0.1599, indicating that there were differences between probability calibration and classification performance [24]. The framework is overall a combination of ordinal prediction, awareness assessment, explainability, cross-IPO validation, and calibration. Yet, there are still some limitations in sampling, discrimination and external validation. Larger independent samples and prospective longitudinal behavioural data should be included in future studies to enhance generalizability and reliability of decisions for support.

6. Conclusion

In this study, explainable ordinal machine-learning models AE-OXGB v4 and AE-OCB were designed and tested for predicting IPO subscription decisions of retail investors based on demographic, behavioural and awareness factors. A total of 1,936 respondent–IPO observations were analysed from 11 IPOs and 176 UAE retail investors. This framework incorporated the following components: respondent-grouped nested cross validation, awareness ablation, subgroup analysis, cross-IPO validation, permutation feature importance, and probability calibration, that were used to analyse multiple aspects of predictive reliability.
For overall classification discrimination, although AE-OCB was lower (0.9029) than AE-OXGB v4 (0.9520), there was not a significant difference between the two. Both models showed prediction errors were reduced by awareness-feature exclusion, indicating that supplemental informational variables were not consistently beneficial for prediction. Effect sizes were found to be small and unstable across investor profiles, and the prediction errors were found to vary across the different investor profiles, suggesting that they are heterogeneous. Similarly, when comparing the offerings across the cross-IPD, there were differences in the predictive transferability of the offerings.
Results showed that temperature scaling reduced the logarithmic loss from 1.5508 to 1.4068 and decreased the Brier score, while macro-F1 decreased, showing the difference between the reliability of probabilistic forecasts and categorical discrimination. These findings set the groundwork for the need to evaluate separately the informational relevance, incremental predictive usefulness, and reliability for decisions.
The study provides not a new learning algorithm but an integrated computational approach, to ordinal financial preference prediction. However, there is some predictive discrimination, small sample size and no external independent validation. Future research should include traditional ordinal baselines, larger investor data sets, longitudinal behavioural data and prospective validation to increase the robustness, interpretability, and relevance for financial decision support.

References

1. L. X. Liu, R. Lu, A. E. Sherman, and Y. Zhang, “IPO underpricing and limited attention: Theory and evidence,” Journal of Banking & Finance, vol. 154, Art. no. 106932, 2023, doi: 10.1016/j.jbankfin.2023.106932.
2. Y. Han and C. Zhang, “Social media analysts, investor attention, and IPO underpricing,” The British Accounting Review, Art. no. 101748, 2025, doi: 10.1016/j.bar.2025.101748.
3. S. Mengoli and P. Pattitoni, “Spotlight on the neighborhood: The spillover effect of IPOs on retail investor attention,” Journal of Empirical Finance, vol. 88, Art. no. 101752, 2026, doi: 10.1016/j.jempfin.2026.101752.
4. D. Pirayesh Neghab, R. Bradrania, and R. Elliott, “Deliberate premarket underpricing: New evidence on IPO pricing using machine learning,” International Review of Economics & Finance, vol. 88, pp. 902–927, 2023, doi: 10.1016/j.iref.2023.07.008.
5. S. Warkulat and M. Pelster, “Social media attention and retail investor behavior: Evidence from r/wallstreetbets,” International Review of Financial Analysis, vol. 96, Art. no. 103721, 2024, doi:

10.1016/j.irfa.2024.103721.
6. M. Münster, F. Reichenbach, and M. Walther, “Robinhood, Reddit, and the news: The impact of traditional and social media on retail investor trading,” Journal of Financial Markets, vol. 71, Art. no. 100929, 2024, doi: 10.1016/j.finmar.2024.100929.
7. J. Xu, M. Khraiche, X. Mao, and X. Wang, “Hubris or talent? Estimating the role of overconfidence in Chinese households' investment decisions,” International Review of Financial Analysis, vol. 91, Art. no. 102992, 2024, doi: 10.1016/j.irfa.2023.102992.
8. E. R. Lawrence, T. D. Nguyen, and B. Wick, “Gender difference in overconfidence and household financial literacy,” Journal of Banking & Finance, Art. no. 107237, 2024, doi: 10.1016/j.jbankfin.2024.107237.
9. P. Weber, K. V. Carl, and O. Hinz, “Applications of explainable artificial intelligence in finance: A systematic review of finance, information systems, and computer science literature,” Management Review Quarterly, vol. 74, pp. 867–907, 2024, doi: 10.1007/s11301-023-00320-0.
10. J. Černevičienė and A. Kabašinskas, “Explainable artificial intelligence (XAI) in finance: A systematic literature review,” Artificial Intelligence Review, vol. 57, Art. no. 216, 2024, doi: 10.1007/s10462-024-10854-8.
11. X. Wu and B. Zhang, “Retail investors' escaping from the bottom and clustering at the top of the trend in China,” International Review of Economics & Finance, vol. 93, pp. 875–904, 2024, doi: 10.1016/j.iref.2024.04.032.
12. H. Aman, T. Motonishi, K. Ogawa, and K. Omori, “The effect of financial literacy on long-term recognition and short-term trade in mutual funds: Evidence from Japan,” International Review of Economics & Finance, vol. 89, pp. 762–783, 2024, doi: 10.1016/j.iref.2023.10.006.
13. A. Oehler, M. Horn, and S. Wendt, “Investment in risky assets and participation in the financial market: Does financial literacy matter?,” International Review of Economics, vol. 71, pp. 19–45, 2024, doi: 10.1007/s12232-023-00432-9.
14. M. Sajid, R. Mushtaq, G. Murtaza, D. Yahiaoui, and V. Pereira, “Financial literacy, confidence and well-being: The mediating role of financial behavior,” Journal of Business Research, vol. 182, Art. no. 114791, 2024, doi: 10.1016/j.jbusres.2024.114791.
15. T. U. Kuzubaş and B. Saltoğlu, “Survey-based measures of risk attitudes and portfolio risk: Evidence from pension participants,” Journal of Behavioral and Experimental Finance, vol. 43, Art. no. 100973, 2024, doi: 10.1016/j.jbef.2024.100973.
16. Y. Chen, R. Calabrese, and B. Martin-Barragan, “Interpretable machine learning for imbalanced credit scoring datasets,” European Journal of Operational Research, vol. 312, no. 1, pp. 357–372, 2024, doi: 10.1016/j.ejor.2023.06.036.
17. Y. Liu, F. Huang, L. Ma, Q. Zeng, and J. Shi, “Credit scoring prediction leveraging interpretable ensemble learning,” Journal of Forecasting, vol. 43, no. 2, pp. 286–308, 2024, doi: 10.1002/for.3033.
18. M. K. Nallakaruppan, B. Balusamy, M. Lawanya Shri, V. Malathi, and S. Bhattacharyya, “An explainable AI framework for credit evaluation and analysis,” Applied Soft Computing, vol. 153, Art. no. 111307, 2024, doi: 10.1016/j.asoc.2024.111307.
19. T. Klein and T. Walther, “Advances in explainable artificial intelligence (XAI) in finance,” Finance Research Letters, vol. 70, Art. no. 106358, 2024, doi: 10.1016/j.frl.2024.106358.
20. P. Buczak, D. Horn, and M. Pauly, “Old but gold or new and shiny? Comparing tree ensembles for ordinal prediction with a classic parametric approach,” Journal of Classification, vol. 42, pp. 364–390, 2025, doi: 10.1007/s00357-024-09497-9.
21. G. Binotto and R. Delgado, “Adapting performance metrics for ordinal classification to interval scale: Length matters,” Machine Learning, vol. 114, Art. no. 41, 2025, doi: 10.1007/s10994-024-06654-4.
22. S. Haas and E. Hüllermeier, “Aleatoric and epistemic uncertainty measures for ordinal classification through binary reduction,” Machine Learning, vol. 115, Art. no. 63, 2026, doi: 10.1007/s10994-025-06960-5.
23. R. Delgado, “Ordinal classification with label-dependent loss,” Machine Learning, vol. 115, Art. no. 60, 2026, doi: 10.1007/s10994-026-07023-z.
24. T. Silva Filho et al., “Classifier calibration: A survey on how to assess and improve predicted class probabilities,” Machine Learning, vol. 112, pp. 3211–3260, 2023, doi: 10.1007/s10994-023-06336-7.
25. P.-H. Huang, “Residual permutation tests for feature importance in machine learning,” British Journal of Mathematical and Statistical Psychology, published online Aug. 2025, doi: 10.1111/bmsp.70009.
26. I. Niankara, “Data for the determinants of IPO subscription among domestic retail investors in the United Arab Emirates,” Mendeley Data, ver. 1, 2022, doi: 10.17632/647nbnm5ss.1.