Handling Multicollinearity in Credit Risk Modelling through Regularisation: an AIC-Guided Elastic Net Approach

Nilton Cardoso
Nick Hewes
Kamakshi Bansal
Monalisa Sinha
Rob Fitton
Scarlett Angel

August 15, 2025

Abstract

Could information criteria guide elastic net regularisation in credit risk? Empirical results using simulation over artificially generated multicollinearity in credit risk models.

Multicollinearity frequently complicates financial modelling, particularly in credit risk assessment, leading to unstable coefficient estimates and potentially overfitting. While Elastic Net regularisation, encompassing L1 (Lasso) and L2 (Ridge) penalties, effectively addresses this, selecting the optimal balance between these penalties remains a challenge. This research empirically tests the utility of Information Criteria (IC) in guiding the selection of Elastic Net parameters (α and λ) when faced with varying levels of multicollinearity.

Our methodology employs a controlled simulation approach. We start with a well-calibrated credit risk dataset exhibiting good predictive performance and minimal overfitting, serving as our benchmark 'ideal model'. We then systematically introduce multicollinearity, first by adding very similar characteristics already existing in the database which are highly correlated with some of the original variables. Then, in a second experiment, we generate some synthetic new features highly correlated (e.g., 90%, 80%, 70% overlap) with the original predictors using techniques like adding controlled noise or linear combinations. Elastic Net models are trained on these modified datasets. We employ various IC metrics – including AIC, AICc, and BIC – to guide the selection of Elastic Net parameters, evaluating how well these selections approximate the benchmark model's performance.

Performance is assessed using standard credit risk metrics (e.g., AUC-ROC) and statistical measures (e.g., R^2, coefficient stability). By comparing the IC-guided regularised models against the benchmark in this controlled environment, we analyse how effectively different criteria manage the bias-variance trade-off under induced multicollinearity. The study aims to provide some practical guidance for practitioners on leveraging IC to tune Elastic Net models robustly, enhancing model stability and reliability when dealing with common real-world data challenges like correlated predictors.

Keywords: Regularisation, Multicollinearity, Elastic Net, AIC, Credit Risk, Modelling

1 Introduction

In the financial sector, particularly in credit risk modelling, multicollinearity is a pervasive issue characterised by high correlations among predictor variables. When predictors are highly correlated, isolating the unique contribution of each variable to credit risk becomes challenging (Daoud, 2017). This can lead to unreliable statistical estimates, where the standard errors of regression coefficients inflate, making them unstable (Midi et al., 2010; Upendra and Teli, 2023). Such instability implies that even small changes in data can drastically alter model coefficients on their standard deviation, resulting in potential inconsistency or misleading interpretations (Kim, 2019).

1.1 Regularisation Techniques: A Remedy for Multicollinearity

Traditional techniques like stepwise regression or principal component analysis (PCA) offer partial solutions but lack interpretability or efficiency in high-dimensional settings. In contrast, regularisation techniques provide robust alternatives by adding penalty terms to the loss function to control model complexity and prevent overfitting (Hajihosseinlou et al., 2024).

1.1.1 Lasso Regression (L1 Penalty)

Lasso, introduced by Tibshirani (Tibshirani, 1996), penalises the absolute values of coefficients:

$$\hat{\beta}_{\text{lasso}} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum (y_i - X_i\beta)^2 + \lambda \sum |\beta_j| \right\} \tag{1}$$

Lasso induces sparsity by shrinking some coefficients exactly to zero, effectively selecting a subset of features. However, it tends to select only one variable from a group of highly correlated predictors, potentially causing instability.

1.1.2 Ridge Regression (L2 Penalty)

Ridge regression, introduced by Hoerl and Kennard Hoerl and Kennard (2000), shrinks the regression coefficients by adding a penalty proportional to their squared magnitude:

$$\hat{\beta}_{\text{ridge}} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum (y_i - X_i\beta)^2 + \lambda \sum \beta_j^2 \right\} \tag{2}$$

While Ridge effectively reduces multicollinearity, it does not perform feature selection as it keeps all predictors in the model.

1.1.3 Elastic Net (L1 + L2 Penalty)

To combine the benefits of both Ridge and Lasso, Zou and Hastie (Zou and Hastie, 2005) introduced Elastic Net. It applies a convex combination of both penalties:

$$\hat{\beta}_{\text{EN}} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum (y_i - X_i\beta)^2 + \lambda \left[\alpha \sum |\beta_j| + (1-\alpha) \sum \beta_j^2 \right] \right\} \tag{3}$$

Here, λ controls the overall strength of regularisation, and $\alpha \in [0, 1]$ determines the balance between Lasso (L1) and Ridge (L2). Elastic Net not only handles multicollinearity but also introduces grouping effects, stabilising the model and improving interpretability (Kerata and Naumova, 2022). Elastic Net has gained traction in various sectors, including credit scoring and stroke prediction, for its effectiveness in identifying influential features and enhancing model performance (Merdas, 2024).

1.2 AIC as a Model Selection Tool for Elastic Net

While regularisation methods are powerful, tuning the regularisation parameters λ and α is a non-trivial task. Cross-validation (CV) is widely used but becomes computationally expensive and unstable in high-dimensional settings. An alternative is the Akaike Information Criterion (AIC), introduced by Akaike (Akaike, 1973), and Bayesian information criterion (BIC) (Schwarz, 1978) are key model selection tools in econometrics, particularly for forecasting. Both criteria estimate out of sample forecast error variance by penalising in-sample residual mean squared error to account for model complexity (degrees of freedom) providing a more robust approach Diebold (2016).

Information Criterion Formulas:

- AIC: AIC = -2 ln($\hat{L}$) + 2K
- BIC: BIC = -2 ln($\hat{L}$) + K ln(N)

Where $\hat{L}$ is the maximised log-likelihood, K is the number of model parameters, and N is the number of observations. AIC generally favors more complex models, while BIC penalises complexity more harshly, especially in large samples Diebold (2016). While some research suggests standard AIC/BIC may be infeasible for moderately large models Chen et al. (2018), other studies have demonstrated its effectiveness. For instance, Jordanger and Tjøstheim (Jordanger and Tjøstheim, 2014) compared penalised AIC (pAIC) with a cross-validation-based Copula Information Criterion and found that both perform similarly for large datasets. Kamel et al. (Kamel et al., 2017) also showed that AIC outperforms RMSE and R^2 in model evaluation tasks. This suggests that AIC may offer a reliable, computationally efficient alternative to CV for parameter tuning in Elastic Net.

In credit risk modelling, where multicollinearity is a significant concern, Elastic Net offers a promising regularisation strategy that combines the strengths of both Ridge and Lasso. However, effective model performance hinges on tuning the regularisation parameters, and AIC-based approaches provide a compelling alternative to traditional cross-validation, especially in high-dimensional contexts. This research explores and advocates the integration of AIC with Elastic Net for efficient and interpretable credit risk modelling.

2 Dataset

A benchmark predictive model was built using a comprehensive, anonymized sample of 639,612 unsecured loan accounts from the Equifax database that were originated between January 2020 and July 2022, with repayment terms from one to five years. For this analysis, "bad performance" was defined as three or more payments in arrears within a twelve-month period after the loan's origination. All other performance was categorized as "good". The sample included 97% good accounts and 3% bad accounts, with 1,009 distinct variables per observation. The Information Value (IV) was used to identify variables with strong predictive power.

Candidate variables were manually coarse-classed to show a logical trend with the target variable (1 = good, 0 = bad) based on their Weight of Evidence (WoE) and bad rate. After finalising these coarse classes, a new WoE variable was created for each, based on the formula:

$$\text{WoE}_i = \text{Ln}\left(\frac{\%\text{Goods}_i}{\%\text{Bads}_i}\right). \tag{4}$$

A baseline logistic regression model was then built using these selected WoE variables to estimate the probability of a loan account being "good".

Data Summary

- **Application Date (Start_Date) Range:** 1st January 2020 to 31st July 2022
- **Loan Type:** Unsecured Personal Loans Only (Account_Type=2)
- **Loan Term:** 1 to 5 years
- **Performance Definition:** Bad is 3 payments down in 12 months otherwise Good.
- **Characteristics:** Only First Applicant Current Address characteristics included along with the Risk Navigator scores - RNILF05, RNILF04, RVILF04, FTILF04 and MGILF04.
- **Observations and Variables:** 639,612 observations and 1,009 variables.

Sample Split

This sample has split as follows.

- **Development Sample (70%):** Date Range: 1st January 2020 to 31st December 2021, Bads: 8,127 (2.49%), Goods: 317,958 (97.51%)
- **Hold-Out Sample (30%):** Date Range: 1st January 2020 to 31st December 2021, Bads: 3,638 (2.61%), Goods: 135,693 (97.39%)
- **Out of Time Sample:** Date Range: 1st January 2022 to 31st July 2022, Bads: 5,053 (2.90%), Goods: 169,143 (97.10%)

Methodology

The experimental phase of this research aimed to test whether an AIC-based approach could effectively tune Elastic Net's regularization parameters to mitigate multicollinearity.

First, a benchmark predictive model was established using a comprehensive sample of 639,612 anonymized unsecured loan accounts. These loans were originated between January 2020 and July 2022 with repayment terms of one to five years. For this analysis, "bad performance" was strictly defined as three or more payments in arrears within twelve months of origination, while all other performance was classified as "good." The sample reflected a 97% "good" and 3% "bad" performance split. The dataset for the benchmark model included 1,009 distinct variables per observation. The Information Value (IV) was used to identify variables with strong predictive power. Candidate variables were manually coarse-classed to ensure a logical trend with the target variable, and a new variable was created using the Weight of Evidence (WoE) for each coarse class.

Experiment number 1 used four actual predictor variables that were highly correlated with existing baseline variables. These variables were chosen based on their logical relationship to existing predictors, often representing similar information over a different time period. The presence of severe multicollinearity was confirmed through diagnostic testing; the correlation matrix showed a correlation of 0.9 or greater between all candidate and baseline variables. Variance Inflation Factors (VIFs) and Tolerances further confirmed that including these variables would introduce problematic multicollinearity.

On the experiment number 2, the correlated features of the experiment 1 were replaced with six new, highly-correlated synthetic features (COR001 to COR006). These variables were created as a combination of existing baseline features with added random noise to simulate real-world data variability.

The newly created dataset, which included both the original baseline predictors and the identified correlated variables, served as the testbed. The Elastic Net model's performance, with its parameters tuned via AIC, was evaluated for model stability, variable selection consistency, and predictive accuracy (Gini coefficient). The results were then compared to the non-regularized benchmark model using its pre-calculated AIC and BIC statistics.

The dataset itself featured 1,009 distinct variables per observation and the Information Value (IV) was used to identify variables that have strong predictive power to discriminate between “good” and “bad” accounts.

Candidate variables were manually coarse classed so there is a logical trend with the target variable (1 = ”Good”, 0 = “Bad”) with regards to the Weight of Evidence (WoE) and bad rate. Once the coarse classes were finalised a new variable was created that consisted of the WoE for the corresponding coarse class.

5

The WoE formula for a given coarse class i is:

$$\text{WoE}_i = \ln\left(\frac{\%\text{Goods}_i}{\%\text{Bads}_i}\right)$$

As an example below are the coarse classes corresponding to the variable number of credit searches in the last 12 months.

Credit_Seeking (searches)	**Number of Goods**	**Number of Bads**	**% Goods**	**% Bads**
0-1	200	2	0.05	0.01
2-5	500	10	0.13	0.05
6-10	1,000	30	0.25	0.15
> 10	2,300	158	0.57	0.79

Table 1: Example of Coarse Classes and WoE Calculation Data

From the original `Credit_Seeking` variable, a WoE variable was created using the WoE values in the table above. So if the observation in the sample had 1 credit search, the corresponding value of the WoE variable is 0.0197204 and so on. The baseline model was then built using logistic regression and is summarized in the table below.

Table 2: Baseline Model Summary

Variable	**Variable Description**	**Parameter Estimate**	**Highest Correlation**	**VIF**	**Tolerance**
Intercept		3.665540			
MM24	Repayment_History	0.536244	0.39 (CC12)	1.247113	0.801852
CC12	Debt_Management	0.322089	0.39 (MM24)	1.345546	0.743193
ZA24	Arrears_Status	0.309260	0.52 (NM4+)	1.454596	0.687476
CS12	Credit_Seeking	0.676157	0.19 (NM4+)	1.059713	0.943651
CAAge	Account_Longevity	0.325705	0.33 (CU1)	1.273968	0.784949
LR	Residence_Stability	0.419904	0.28 (CAAge)	1.121345	0.891786
NM4+	Capacity_arrears_M4plus	0.266096	0.52 (ZA24)	1.447383	0.690902
CU1	Capacity_Utilisation_recent	0.323913	0.33 (CAAge)	1.198951	0.834062

3 Experimentation

3.1 Baseline Model

The foundational model, a logistic regression constructed with eight features (MM24, CC12, ZA24, CS12, CAAge, NM4+, CU1, LR), served as the benchmark for subsequent experimental analyses. Its predictive performance was notably robust, evidenced by an AUC of 0.8363 on the training sample and 0.8177 on the Out-of-Time (OOT) sample. In terms of model selection, its

6

information criteria values were determined to be AIC = 63001.68 and BIC = 63097.93; lower values for these criteria signify an enhanced equilibrium between model fit and complexity.

3.2 Experiment 1: Correlated characteristics existing in the original dataset

3.2.1 Model Updates

In this experiment, two new features, MM12 and NM3+, were introduced with observed correlation with existing, established baseline features. Specifically, MM12 demonstrated a notable correlation with MM24, whilst NM3+ was found to be correlated with NM4+. This inherent correlation necessitated careful consideration during the experimental design and analysis, as it contributed to the creation of what is termed an "inflated" data set. The "inflated" data set is characterised by its expanded dimensionality, now comprising a total of 10 distinct features.

Initially, a non-regularised logistic regression model was fitted on the inflated dataset. Following this, to address the introduced multicollinearity, various regularisation techniques were applied. These techniques, such as Ridge, Lasso, and Elastic-Net regression, work by adding a penalty term to the logistic regression's cost function, effectively shrinking the coefficient values and potentially setting some to zero, thereby mitigating the effects of multicollinearity and preventing overfitting. The application of these regularisation methods allowed for a more robust and interpretable model, better able to handle the complexities of the inflated dataset.

To pinpoint the most robust and accurate predictive model, a comprehensive comparative analysis was undertaken, involving the fitting of several distinct model specifications. A systematic approach was adopted whereby, in each iteration, a different highly correlated feature was strategically removed from the dataset. This iterative process aimed to mitigate multicollinearity and assess the impact of feature exclusion on model performance.

Following the initial model fitting, a meticulous hyperparameter optimisation process was performed for each model. This involved fine-tuning the critical hyperparameters, specifically regularisation strength (λ) and elastic net mixing parameter (α). To ensure a thorough and unbiased optimisation, a variety of objective functions were employed. These included: AUC optimization, negative log-loss optimization, and direct minimization of AIC and BIC.

The regularisation methods aimed to select a more robust and interpretable model, handling the complexities of the inflated dataset towards a more appropriate model.

3.2.2 Results from Experiment 1

The findings indicated that Elastic Net models optimized by AIC and BIC consistently achieved the lowest information criteria values. For example, one with none of the chars dropped had an AIC of 62966.53, which is lower than the baseline's AIC of 63001.68. A theoretically

7

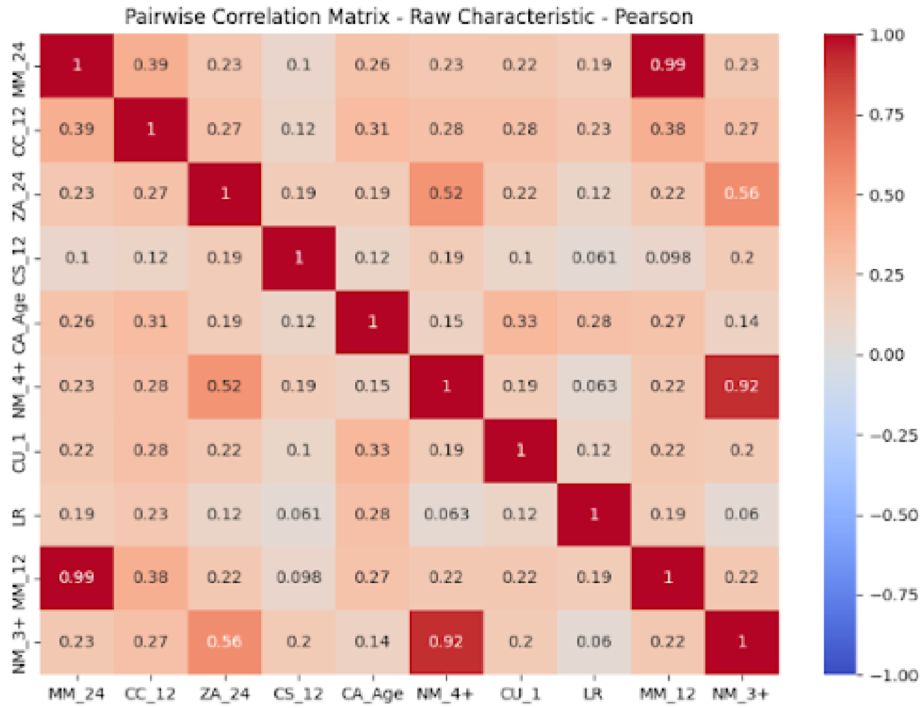

Figure 1: Correlation matrix for variables used in experiment 1

inconsistent finding was that the Inflated model's AIC (62966.54) was nearly identical to that of the best Elastic Net models. This is unexpected because the Inflated model has more features and no explicit regularization, yet it performed just as well in terms of AIC/BIC. This suggests a specific characteristic of the dataset where the additional correlated features did not significantly penalize the AIC.

Another surprising result is that the IC for the baseline model surpasses all other models, which is not within the expectation.

3.3 Experiment 2: Synthetic Characteristics, artificially generated

3.4 Model Updates

In this experiment, the correlated features were replaced with six new, highly-correlated synthetic features (COR001 to COR006). These variables were created as a combination of existing baseline features with added random noise to simulate real-world data variability. With these additions, the dataset now contains a total of 14 features (8 original and 6 synthetic).

Table 3: Experiment 1 Results

Removed Variable	Model Type	Training Sample		OOT Sample		
		AIC	BIC	Accuracy	AUC	Precision
None	Baseline (No Corr, No reg)	63001.68	63097.93	97.10%	81.77%	97.10%
	Inflated (No Reg.)	62966.54	63084.18	97.10%	81.79%	97.10%
	L1 (Log Loss Opt)	62966.53	63084.17	97.10%	81.79%	97.10%
	L2 (Log Loss Opt)	62966.53	63084.17	97.10%	81.79%	97.10%
	ElasticNet (Log Loss Opt)	62966.60	63084.24	97.10%	81.79%	97.10%
	ElasticNet (AIC Opt)	62966.53	63084.17	97.10%	81.79%	97.10%
	ElasticNet (BIC Opt)	62966.53	63084.17	97.10%	81.79%	97.10%

The synthetic features are defined as follows:

$$\text{COR001} = (0.1 \times \text{CAAge}) + (0.9 \times \text{MM24}) + N(\mu = 0, \sigma = 0.5)$$
$$\text{COR002} = (0.8 \times \text{CC12}) + (0.9 \times \text{CU1}) + (0.95 \times \text{CS12}) + N(\mu = 0, \sigma = 0.5)$$
$$\text{COR003} = \text{LR} + N(\mu = 0, \sigma = 0.5)$$
$$\text{COR004} = (0.7 \times \text{ZA24}) + (0.9 \times \text{NM4+}) + N(\mu = 0, \sigma = 0.5)$$
$$\text{COR005} = \text{CS12} + N(\mu = 0, \sigma = 0.5)$$
$$\text{COR006} = \text{CAAge} + N(\mu = 0, \sigma = 0.3)$$

This construction ensures a high correlation between the new synthetic features and their primary constituent variables. For example, COR001 is highly correlated with MM24 ($\sim$90%), and COR006 is highly correlated with CAAge ($\sim$94%).

3.4.1 Results from Experiment 2

The analysis of the synthetic characters experiment revealed that the Elastic Net model, dropping COR002 and optimized by AIC & BIC, performed best. It achieved the lowest AIC of 63000.39, an improvement over both the baseline (63001.68) and the Inflated model (63008.35). This demonstrates Elastic Net's ability to perform feature selection by zeroing out redundant or less important features. The model retained the original eight features and the synthetic feature COR006, while the coefficients for COR001 to COR005 were effectively set to zero. Unlike Experiment 1, the AIC and BIC for the best Elastic Net model were clearly superior to the inflated model. This result aligns with theoretical expectations, where regularization helps to

9

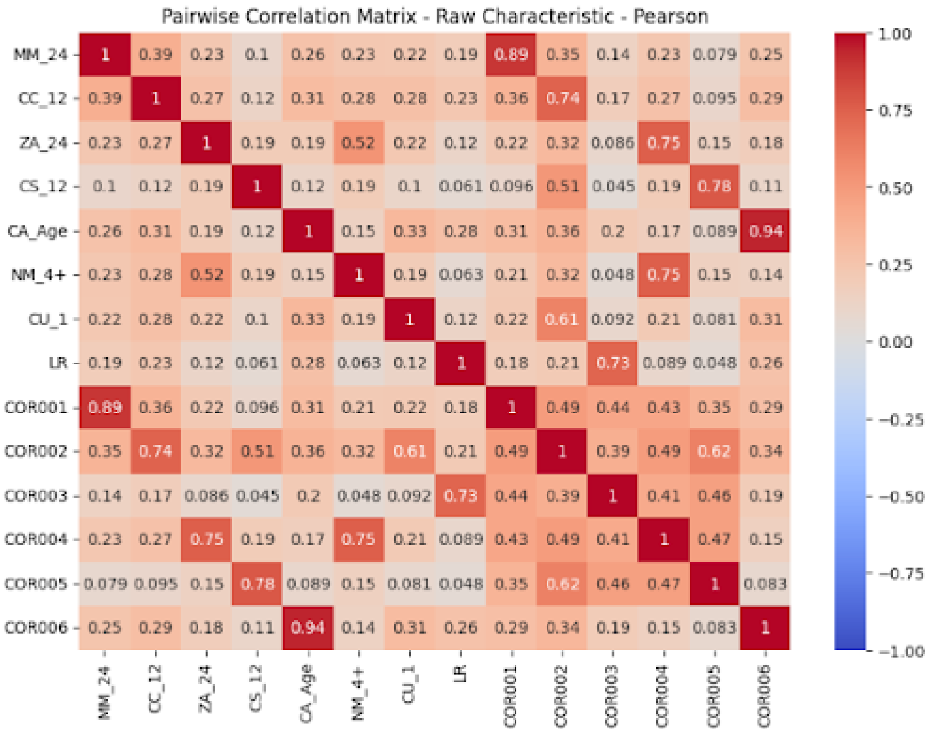

Figure 2: Correlation matrix for variables used in experiment 2

manage the trade-off between model complexity and goodness of fit. The AIC for the base model is quite close to the lowest AIC among all models. The BIC for the base model is the lowest for all the models, within theoretical expectations this would be the best adjusted model.

4 Conclusion

This empirical test aimed to determine the efficacy of using Information Criteria (ICs) to guide Elastic Net regularization in credit risk modeling, especially when tackling multicollinearity. While the initial hypothesis was that the baseline model would consistently show the best fit, as indicated by the lowest AIC and BIC values, the findings presented a more nuanced picture. Only in the second experiment did the baseline model's BIC value prove to be the lowest, reinforcing the theoretical preference for parsimony in that specific scenario. The overall results from the two distinct experiments show that Elastic Net regularization, when guided by ICs, is a valuable method for managing multicollinearity. This approach not only maintains strong predictive performance but also creates a more parsimonious and interpretable model by appropriately shrinking or zeroing out redundant coefficients.

The first experiment, which used a dataset with a few highly correlated real-world variables, yielded an unexpected result: the AIC for the non-regularized "inflated" model was nearly identical to that of the best-performing Elastic Net model. This suggests that when only a limited number of correlated variables are present, they might not introduce enough complexity to be heavily penalized by the ICs. This highlights a dataset-dependent aspect of model selection,

Table 4: Experiment 2 Results

Removed Variable	Model Type	Training Sample		OOT Sample		
		AIC	BIC	Accuracy	AUC	Precision
COR002	Baseline (No Corr, No reg)	63001.68	63097.93	97.39%	83.63%	97.39%
	Inflated (No Reg.)	63008.35	63158.08	97.39%	83.64%	97.39%
	L1 (Log Loss Opt)	63000.57	63107.52	97.39%	83.64%	97.39%
	L2 (Log Loss Opt)	63008.34	63158.07	97.39%	83.64%	97.39%
	ElasticNet (Log Loss Opt)	63000.43	63107.38	97.39%	83.64%	97.39%
	ElasticNet (AIC Opt)	63000.39	63107.34	97.39%	83.64%	97.39%
	ElasticNet (BIC Opt)	63000.39	63107.34	97.39%	83.64%	97.39%

where the theoretical benefits of regularization may not always translate into a significant improvement in IC metrics, particularly when the added correlated features are few.

In contrast, the second experiment, which introduced six distinct synthetic features with varying degrees of controlled multicollinearity, provided a clearer and more theoretically consistent outcome. In this more complex environment, the AIC-optimized Elastic Net model clearly outperformed the non-regularized inflated model, achieving a lower AIC and demonstrating its superior ability to balance model fit with complexity. This experiment strongly supports the central hypothesis: when faced with a more substantial and complex set of correlated predictors, Elastic Net, guided by AIC, effectively performs variable selection and coefficient shrinkage, leading to a demonstrably better model.

The divergence in results between the two experiments underscores a critical lesson for practitioners: the performance and benefit of a regularization technique are highly dependent on the nature and extent of multicollinearity within the data. Although the first experiment served as an important reminder that real-world data can sometimes behave unpredictably, the second experiment provided a robust and reproducible validation of the efficacy of Elastic Net.

In conclusion, while further testing and theoretical investigation are warranted, an AIC-guided Elastic Net approach presents a promising and computationally efficient alternative to traditional methods in credit risk modeling. It offers a viable way to build robust models that are less susceptible to instability from highly correlated predictors, ultimately enhancing the reliability and interpretability of credit risk assessments.

Acknowledgments

We acknowledge the support of Equifax UK for this research. We also thank our internal sponsors and especially Denise Furlong for her valuable help with reference formatting. Any errors or ideas presented are those of the authors alone and do not reflect the position of Equifax.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In Petrov, B. N. and Csáki, F., editors, *Proceedings of the Second International Symposium on Information Theory*, pages 267–281. Akadémiai Kiadó, Budapest.

Chen, X., Aravkin, A. Y., and Martin, R. D. (2018). Generalized linear model for gamma distributed variables via elastic net regularization. *arXiv preprint arXiv:1804.07780*.

Daoud, J. I. (2017). Multicollinearity and regression analysis. *Journal of Physics: Conference Series*, 949:012009.

Diebold, F. X. (2016). Modeling and forecasting trend. In *Elements of forecasting*. Thomson South-Western.

Hajihosseinlou, M., Maghsoudi, A., and Ghezelbash, R. (2024). Regularization in machine learning models for mvt pb-zn prospectivity mapping: applying lasso and elastic-net algorithms. *Earth Sci Inform*, 17:4859–4873.

Hoerl, A. E. and Kennard, R. W. (2000). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 42(1):80–86.

Jordanger, L. A. and Tjøstheim, D. (2014). Model selection of copulas: Aic versus a cross validation copula information criterion. *Statistics & Probability Letters*, 92:249–255.

Kamel, A. H., Shaqlaih, A. S., and Rozyyev, A. (2017). A comparative analysis and evaluation of model selection criteria. In *Proceedings of the ASME 2017 International Mechanical Engineering Congress and Exposition*, volume 5, page V005T06A015. ASME.

Kerata, Z. and Naumova, V. (2022). On an unsupervised method for parameter selection for the elastic net. *Mathematics in Engineering*, 4(6):1–36.

Kim, J. (2019). Multicollinearity and misleading statistical results. *Korean Journal of Anesthesiology*, 72(6):558–568.

Merdas, H. (2024). Elastic net – mlp – smote (ems)-based model for enhancing stroke prediction. *Journal of Medical Informatics and Nursing (MEDIN)*, 1(2):1–8.

Midi, H., Sarkar, S. K., and Rana, S. (2010). Collinearity diagnostics of binary logistic regression model. *Journal of Interdisciplinary Mathematics*, 13(3):253–267.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2):461–464.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288.

Upendra, R. and Teli, S. N. (2023). Multicollinearity in multiple linear regression: Detection, consequences, and remedies. *International Journal for Research in Applied Science and Engineering Technology*, 11(9):1047–1061.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2):301–320.