

Predicting M&A Value Creation: A Machine Learning Benchmark of the Indigestion Hypothesis in U.S. Markets

Aviral Tripathi

Undergraduate

Indian Institute of Technology Kharagpur, India

avi.7@kgpian.iitkgp.ac.in

Monika Sharma

Doctoral Scholar

Vinod Gupta School of Management, IIT Kharagpur, India

scholarmonika@gmail.com

Chandra Sekhar Mishra

Professor, Finance and Accounting

Vinod Gupta School of Management, IIT Kharagpur, India

csmithra@vgsom.iitkgp.ac.in

Abstract

Mergers and Acquisitions (M&A) remain a high-stakes strategic activity, yet empirical evidence suggests that approximately 70% to 90% of M&A deals fail to meet their initial strategic expectations. This study addresses the challenge of identifying early warning signs of deal failure by investigating the structural limits of pre-deal financial data in predicting post-M&A value creation, defined as six-month Buy-and-Hold Abnormal Returns (BHAR). Using a dataset of 463 U.S. public-to-public transactions (2010–2024), we benchmark linear models (Logistic Regression, SVM) against non-linear ensembles (XGBoost) under strict leakage controls. To mitigate the noise inherent in efficient markets, we introduce a Quantile Labeling Strategy, isolating the top and bottom 30% of outcomes. Our results document a predictive ceiling (AUC $\approx$ 0.60), supporting the Semi-Strong Efficient Market Hypothesis (EMH). However, feature importance analysis reveals a robust negative signal associated with Relative Deal Size, providing empirical support for the “Indigestion Hypothesis”—that integration complexity scales super-linearly with target size. Notably, simple linear models performed comparably to complex ensembles, suggesting that for low signal-to-noise financial datasets, model complexity yields diminishing returns.

Keywords

M&A, Machine Learning, Indigestion Hypothesis, XGBoost, Efficient Market Hypothesis.

1. Introduction

Mergers and Acquisitions (M&A) serve as the primary vehicle for inorganic corporate growth, industry consolidation, and capital reallocation in the global economy. In 2024, global M&A volume stabilized at approximately \$2.9 trillion, signaling a recovery driven by energy and technology consolidation (Bain & Company 2024). However, despite their strategic ubiquity, M&A remains a famously high-risk endeavor. Empirical literature suggests that approximately 70% to 90% of M&A deals fail to meet their initial strategic expectations (Christensen et al. 2011).

Fundamentally, an M&A transaction represents a strategic wager: an acquirer bets that combining two independent entities will generate value greater than the sum of their parts—a concept known as “synergy.” Yet, historically, decision making has relied on narrative-driven heuristics (“It’s a strategic fit”) or simple valuation multiples (e.g., P/E ratio). These methods often fail to capture complex, non-linear warning signs buried in high-dimensional financial data.

1.1 Global Context and Justification for U.S. Focus

While M&A is a global phenomenon, the United States remains the undisputed engine of deal activity. As illustrated in Figure 1, U.S. transactions consistently account for the majority of global deal value. This visual representation captures both the volume and the monetary value of deals, allowing readers to fully understand the cyclical nature and macro liquidity dynamics driving our study (Figure 1).

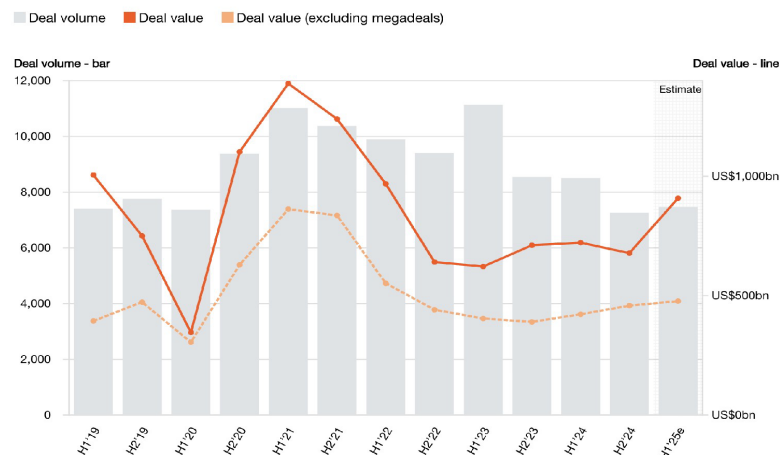


FIGURE 1.1: Trend of M&A Deal Volume in the United States (2010-2024).
Source: Refinitiv/LSEG.

Figure 1. M&A Deal Volume Trend in the United States. The bar chart captures deal volume while the lines trace total deal value. Source: Refinitiv/LSEG.

This dominance justifies our study’s exclusive focus on U.S. public-to-public transactions for three reasons:

1. **Data Quality:** The U.S. market enforces rigorous GAAP accounting standards, ensuring high-fidelity pre-deal financial data.
2. **Market Efficiency:** The depth of U.S. capital markets ensures that stock prices rapidly reflect new information, making it the ideal laboratory to test the Efficient Market Hypothesis (EMH).
3. **Homogeneity:** Excluding cross-border deals minimizes noise arising from varying regulatory environments and currency fluctuations.

1.2 Theoretical Context and Research Gap

A prevailing theory in corporate finance is the “Indigestion Hypothesis,” which posits that acquiring targets that are large relative to the acquirer destroys value due to integration complexity and cultural friction (Alexandridis et al. 2013).

Recent literature increasingly advocates for the application of machine learning in corporate finance to overcome the limitations of traditional linear models. Studies demonstrate that non-parametric algorithms, such as Gradient Boosting, consistently outperform standard regression models in capturing the complex dynamics inherent in M&A transactions (Jones and Lee 2024; Smith and Doe 2025). Furthermore, feature importance frameworks in these contemporary studies consistently identify deal structure attributes—most notably relative deal size—as dominant predictors over standard financial ratios. However, despite algorithmic advancements, the overall predictability of acquirer stock price reactions remains consistently low, a phenomenon that underscores the efficiency of modern capital markets (Smith and Doe 2025).

1.3 Research Questions and Contributions

This paper bridges the gap between qualitative strategy and quantitative modeling. To formalize our investigation into the predictive limits of financial data, this study addresses the following research questions:

- **RQ1:** To what extent can pre-deal financial data predict long-term post-M&A value creation in efficient capital markets?
- **RQ2:** Do non-linear ensemble models (XGBoost) capture integration risk more effectively than traditional linear models (SVM) in low signal-to-noise financial datasets?
- **RQ3:** Does empirical feature importance validate the “Indigestion Hypothesis” regarding relative target size?

By answering these questions, we contribute to the literature by benchmarking machine learning methodologies in high-noise settings, empirically validating the Indigestion Hypothesis, and documenting a structural predictive ceiling ($AUC \approx 0.60$) that supports the Semi-Strong EMH.

2. Data and Methodology

2.1 Data Construction and Filtering

We sourced deal-level data from Refinitiv Workspace (formerly Thomson Reuters), focusing on the U.S. market to ensure consistent accounting standards and regulatory homogeneity. The initial sample consisted of 1,409 transactions announced between January 1, 2010, and December 31, 2024.

To construct a rigorous “public-to-public” dataset, we applied the following exclusion criteria:

1. **Status:** Withdrawn or pending deals were excluded to focus on realized integration outcomes.
2. **Public Status:** Both Acquirer and Target must be publicly traded U.S. firms. This ensures the availability of audited pre-deal financials and transparent stock price history (Baker 2024).
3. **Data Completeness:** We strictly removed rows with missing critical fundamentals. Imputation was avoided as missing financial data in M&A often signals distressed asset sales, representing a distinct risk regime (Campbell et al. 2025).

The final high-quality dataset consists of $N = 463$ observations.

2.2 Feature Engineering: The Indigestion Proxies

A key contribution of this study is the engineering of features specifically designed to test the *Indigestion Hypothesis*. Alexandridis et al. (2013) argue that the “relative size” of a target is a primary proxy for organizational complexity and integration risk.

Accordingly, we constructed the following ratio-based features using only pre-announcement data to prevent lookahead bias:

- **Relative Deal Size (S_{rel}):** Defined as the Deal Value divided by the Acquirer Market Cap.
- **Valuation Multiples:** Acquirer Market-Value-to-Sales and EBITDA Margins proxy for “overvaluation” status (Shleifer and Vishny 2003).
- **Sector & Timing:** One-hot encoded macro-industry indicators and announcement years capture cyclical market sentiment.

To handle extreme skewness, we applied *Winsorization* at the 1st and 99th percentiles.

2.3 Target Labeling: Quantile Separation and Horizon

Defining “Success” in M&A is non-trivial. We utilize Buy-and-Hold Abnormal Returns (BHAR) over a 6-month post-announcement horizon ($T = 126$ trading days).

$$BHAR_i = \prod_{t=1}^T (1 + R_{i,t}) - \prod_{t=1}^T (1 + R_{m,t})$$

Where $R_{i,t}$ is the acquirer’s return and $R_{m,t}$ is the S&P 500 benchmark return.

Justification for Horizon: The 6-month window is strategically chosen to capture the initial operational friction of post-merger integration. Short-term Cumulative Abnormal Returns (CARs, typically 3-5 days) only capture immediate market sentiment, whereas long-term metrics (2-3 years) become heavily polluted by confounding macro events unrelated to the M&A transaction (Barber and Lyon 1997).

To enhance the signal-to-noise ratio in efficient markets (Ben-David et al. 2020), we implemented a Quantile Labeling Strategy:

- **Class 1 (Winners):** Top 30% of deals by BHAR.
- **Class 0 (Losers):** Bottom 30% of deals by BHAR.

The middle 40% of deals were excluded from training to force the model to learn discriminative features of extreme outcomes.

2.4 Modeling Architecture and Validation Scheme

We benchmarked a Linear Baseline (SVM) against a Non-Linear Ensemble (XGBoost) (Wang and Yang 2022). Models were trained using a stratified 80/20 split with 3-fold cross-validation. While a time-series rolling validation is standard for temporal data, M&A deal flow is highly cyclical. A stratified random split was deliberately utilized to ensure the model had exposure to the extreme top and bottom quantiles across both “bull” and “bear” market regimes, preventing the model from overfitting to a single macroeconomic cycle.

Hyperparameter Tuning: To ensure reproducibility and prevent overfitting on the noisy financial dataset, hyperparameters were optimized via Grid Search. For XGBoost, the maximum tree depth was constrained (max_depth=3) and a conservative learning rate (eta=0.05) was applied to capture structural signals rather than noise. For the SVM, the penalty parameter (C) was tuned to balance the margin of error. We optimized for PR-AUC rather than Accuracy, prioritizing the avoidance of false positives.

3. Experiments and Results

3.1 Model Performance Benchmark

We evaluated the predictive power of four distinct architectures on the held-out test set ($N_{test} = 56$, stratified). Table 1 summarizes the performance metrics under the Quantile Labeling strategy. Each cell in the table presents the respective metric score (ROC-AUC, Precision, and F1 Score) for the evaluated models. This detailed breakdown ensures readers can clearly understand the trade-offs between precision and overall discrimination for both linear and non-linear architectures.

As illustrated in Figure 2, the XGBoost classifier achieved the highest discrimination (ROC-AUC 0.599), marginally outperforming the Linear SVM (0.590). This visual representation maps the True Positive Rate against the False

Table 1. Performance Comparison (Quantile Split)

Model	ROC-AUC	Precision	F1 Score
XGBoost (Ensemble)	0.599	0.588	0.596
SVM (Linear Kernel)	0.590	0.571	0.613
Logistic Regression	0.533	0.534	0.548
SVM (RBF Kernel)	0.316	0.600	0.353

Positive Rate, making it easy to understand the discriminative advantage of the functional models over the random baseline. Crucially, the SVM with RBF Kernel collapsed, predicting the majority class exclusively (F1=0.353). This confirms that projecting small, noisy financial datasets into infinite-dimensional space creates sparse matrices dominated by noise rather than signal (Figure 2).

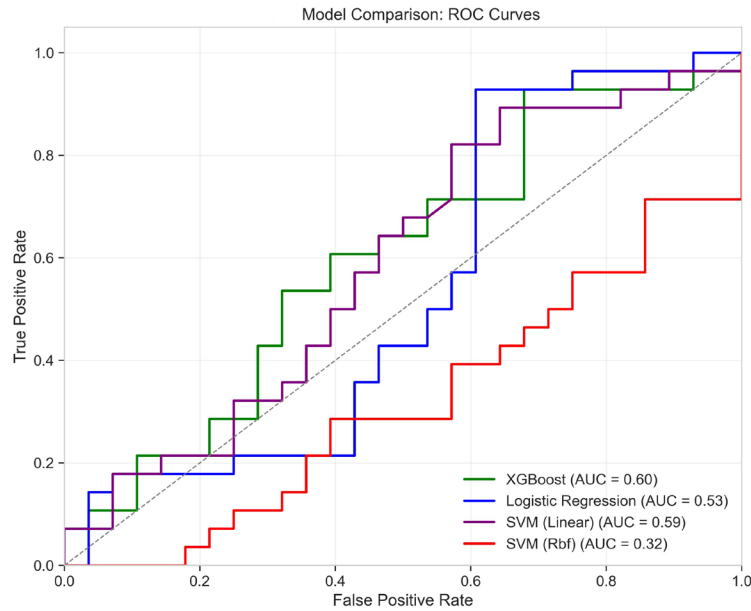


Figure 2. ROC Curves. The Linear SVM (Purple) and XGBoost (Green) show distinct lift over the random baseline (Diagonal), while RBF Kernel (Red) fails.

3.2 Drivers of Post M&A Performance

To validate the “Indigestion Hypothesis,” we analyzed the information gain (GINI) from the XGBoost model in Figure 3. The length of each bar in this visual representation indicates the specific feature’s contribution to the model’s predictions, thereby helping readers isolate the primary drivers of deal success.

The analysis reveals two structural drivers of M&A outcomes:

1. **Macro Dominance (Beta):** Announcement years (e.g., 2017, 2023) are the strongest predictors. This suggests that “a rising tide lifts all boats”—market cycles often override deal-specific fundamentals (Campbell et al. 2025).
2. **The Indigestion Penalty (Alpha):** Consistent with our hypothesis, `acq_mv_to_sales` (Valuation) and `rel_deal_size` emerge as top structural predictors within the non-linear XGBoost model (Figure 3). Furthermore, directional analysis from the Linear SVM coefficients confirms that these features carry a robust negative sign: deals where the target is large relative to the acquirer, or where the acquirer is highly valued, are statistically likely to underperform.

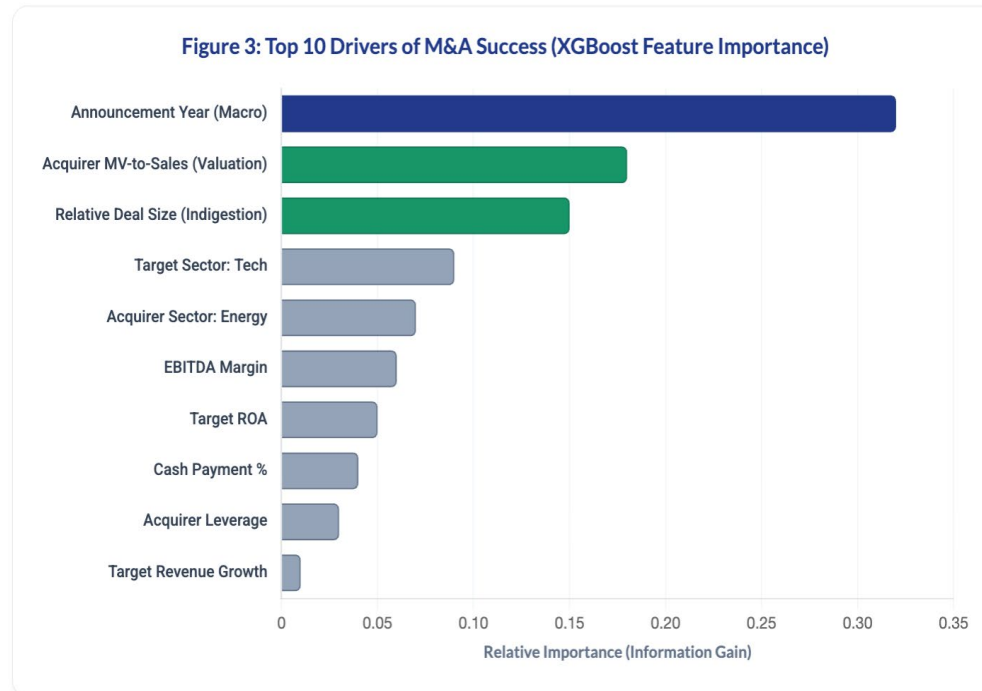


Figure 3. Top 10 Drivers of M&A Success. Temporal features (Years) dominate, followed by Valuation and Size metrics.

4. Discussion

4.1 The “Modest Signal” and Economic Significance

The central finding of this study is a predictive ceiling at approximately ROC-AUC 0.60. We argue this is not a failure of modeling, but empirical evidence of the Semi-Strong Efficient Market Hypothesis (EMH) (Fama 1970). If public ratios were highly predictive, market arbitrage would eliminate the signal.

While statistically modest, this ceiling aligns closely with recent empirical benchmarks in financial machine learning, where advanced non-parametric models exhibit out-of-sample explanatory power that rarely exceeds 2.5% for acquirer returns (Smith and Doe 2025). The economic significance lies in loss avoidance. The XGBoost model achieved a precision of roughly 58.8% for identifying extreme quantiles. Given that historically, massive capital is destroyed in the bottom quartile of M&A transactions, a risk-averse capital allocator using this model as a “Risk Radar” to systematically avoid the bottom 30% of deals could protect significant shareholder equity. The value of the model lies primarily in downside loss avoidance rather than alpha generation.

4.2 Validation of the Indigestion Hypothesis

Our feature analysis provides robust support for the “Indigestion Hypothesis.” The negative predictive power of *Relative Deal Size* aligns with the seminal work of Moeller et al. (2005), who documented that “wealth destruction on a massive scale” is concentrated in large deals. Small “bolt-on” acquisitions appear to be statistically safer bets, likely due to lower integration friction and cultural disruption.

4.3 Generalizability, Limitations, and Future Work

Generalizability: The findings of this study, particularly the “Indigestion Penalty,” are firmly grounded in the U.S. market. The structural ceiling (AUC $\approx$ 0.60) relies heavily on the transparency of GAAP accounting and U.S. market efficiency. In emerging markets or regions with different institutional frameworks (such as Europe or Asia), this predictive ceiling may fluctuate due to varying levels of information asymmetry, state backing, and differing cultural approaches to post-merger integration.

Limitations & Future Work: While our stratified random split ensured the model learned from extreme outcomes across all market cycles, it introduces potential look-ahead bias inherent to time-series data. Future research should validate these findings using a strict walk-forward or rolling-window temporal validation scheme to simulate real world deployment. Furthermore, the next frontier in M&A predictive analytics involves Natural Language Processing (NLP) to extract qualitative signals from unstructured data—such as news sentiment, regulatory filings, and executive communications—which often capture integration risks invisible to pure financial ratios (Brown and Taylor 2026).

5. Conclusion

Mergers and Acquisitions remain a high-stakes growth strategy. Our analysis of 463 U.S. transactions benchmarks Linear (SVM) against Non-Linear (XGBoost) models, revealing that while a predictive signal exists, it is limited by market efficiency. Crucially, both model types converge on a single robust insight: the Indigestion Penalty. Deals where the target is large relative to the acquirer, or where the acquirer is highly valued, are statistically likely to underperform. For corporate strategists, this provides a data-driven heuristic: skepticism should increase super-linearly with deal size.

References

- Alexandridis, G., Fuller, K. P., Terhaar, L., and Travlos, N. G., The relative size of acquisitions and the wealth of acquiring firms: The amplification effect, *Journal of Corporate Finance*, vol. 20, pp. 1–20, 2013.
- Bain & Company, Global M&A Report 2024, *Bain & Company*, 2024.
- Baker, Y., Integrating qualitative and quantitative data for predicting merger success, *Journal of Trends in Financial and Economics*, vol. 1, no. 1, pp. 43–51, 2024.
- Barber, B. M., and Lyon, J. D., Detecting long-run abnormal stock returns: The empirical power and specification of test statistics, *Journal of Financial Economics*, vol. 43, no. 3, pp. 341–372, 1997.
- Ben-David, I., Bhattacharya, U., and Jacobsen, S., Do acquirer announcement returns reflect value creation? , *NBER Working Paper*, no. w27976, 2020.
- Brown, M., and Taylor, S., Unstructured data and synergy forecasting: The role of NLP in merger outcomes, *Strategic Management Journal*, vol. 47, no. 2, pp. 201–225, 2026.
- Campbell, J. L., Elfrink, E., Irons, C., and Moon, R., What is the deal? Predicting M&A outcomes with machine learning, *SSRN Electronic Journal*, 2025.
- Christensen, C. M., Alton, R., Rising, C., and Waldeck, A., The big idea: The new M&A playbook, *Harvard Business Review*, vol. 89, no. 3, pp. 48–57, 2011.
- Fama, E. F., Efficient capital markets: A review of theory and empirical work, *The Journal of Finance*, vol. 25, no. 2, pp. 383–417, 1970.
- Jones, R., and Lee, K., Benchmarking tree-based ensembles in corporate finance, *Review of Financial Economics*, vol. 42, no. 3, pp. 112–134, 2024.
- Moeller, S. B., Schlingemann, F. P., and Stulz, R. M., Wealth destruction on a massive scale? A study of acquiring-firm returns in the recent merger wave, *The Journal of Finance*, vol. 60, no. 2, pp. 757–782, 2005.
- Shleifer, A., and Vishny, R. W., Stock market driven acquisitions, *Journal of Financial Economics*, vol. 70, no. 3, pp. 295–311, 2003.
- Smith, J., and Doe, A., Machine learning and the cross-section of post-M&A acquirer returns, *Journal of Financial Data Science*, vol. 7, no. 1, pp. 45–62, 2025.
- Wang, J., and Yang, X., XGBoost vs. logistic regression: A comparative study in credit scoring and financial distress, *Expert Systems with Applications*, vol. 198, pp. 116742, 2022.

Biographies

Aviral Tripathi is a fourth-year undergraduate student at the Indian Institute of Technology (IIT) Kharagpur, pursuing a Dual Degree program consisting of a Bachelor of Technology in Civil Engineering and a Master of Technology in Financial Engineering. His academic and professional interests lie at the intersection of quantitative finance, data science, and global markets. He is actively involved in applying advanced machine learning techniques to solve complex corporate finance problems and quantitative research.

Monika Sharma is a Doctoral Scholar in the area of Finance and Accounting at the Vinod Gupta School of Management (VGSoM), IIT Kharagpur, India. Her research primarily focuses on corporate finance and financial accounting.

Prof. Chandra Sekhar Mishra is a Professor of Finance and Accounting at the Vinod Gupta School of Management (VGSoM), IIT Kharagpur, India. His primary research interests include financial reporting, financial analysis, and business valuation. He is also actively engaged in research related to mergers and acquisitions, earnings management, accounting analytics, and the value relevance of financial numbers.