Private Equity M&A Target Prediction with Legal and Risk-Aware Machine Learning

Translating Market Signals into Buyout Logic through Predictive Modeling

Calvin Lee | Apr 2025

Analysis Overview

Objective:

- Identify potential PE acquisition targets in public markets
- Combine structured financial filters + legal complexity + predictive scoring (Machine Learning)

Data Source:

- S&P 500 constituents
- Real-time financials via YFinance library

Part I – S&P 500 Screening

Manual rule-based filtering applied across the S&P 500 universe

Reflects strict acquisition criteria used by private equity firms over the past five years (Post-pandemic)

Focused on realistic buyout potential considering size, profitability, and capital structure

S&P 500 Filter Criteria

Market Cap: $1B to $30B
→ Typical size range targeted by PE for control deals

EBITDA Margin ≥ 10%
→ Minimum profitability level for sustainable buyout

ROA > 0
→ Signals productive use of capital

Debt/Equity < 4.0
→ Avoids firms with excessive leverage

Sectors: 7 major M&A-active industries
→ Covers 80~85% of PE deals in past five years
(Technology / Healthcare / Industrials / Real Estate / Consumer Discretionary / Financials / Energy)

Part I - Result

Final candidates after filtering - ERIE, TROW, TPL

Only three companies out of the S&P 500 passed all criteria

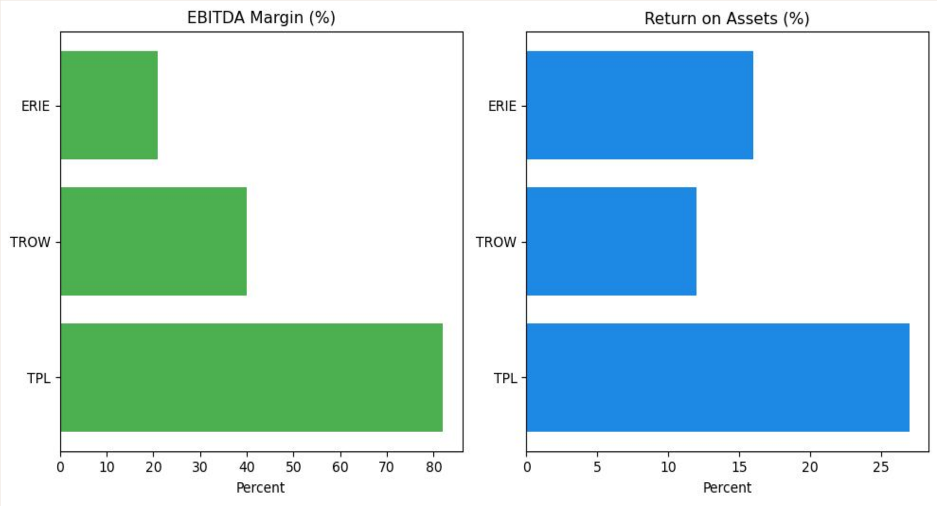
Most were excluded due to excessive size, low profitability or high leverage

The outcome confirms that the S&P 500 is not a realistic field for typical PE control deals

Financial Comparison

TPL shows strong margin and asset returns fit for PE structuring

TROW and ERIE meet thresholds but show lower efficiency and profitability



Buyout Feasibility Analysis

- ERIE (Erie Indemnity)

Operates in a regulated insurance sector with restrictive ownership and required approval for control change

- TROW (T Rowe Price)

Asset manager reliant on brand and licensing structure with limited flexibility for acquisition

- TPL (Texas Pacific Land)

Stable cash flow from land royalties with a clean structure and minimal legal friction

Conclusion

Only **TPL** presents a realistic opportunity for acquisition. The others face legal and structural limitations

Final Candidate Profile (S&P 500)

Texas Pacific Land Corporation (TPL)

Owns large land holdings in Texas with revenue from royalties, easements and leases

Simple operations with high margin and recurring cash flow

Minimal debt with strong scalability and cash generation

Well suited for private equity through income growth and asset monetization

Part II - NASDAQ Mid-Cap Expansion

To expand the scope, the same criteria were applied to NASDAQ-listed MidCap companies.

Thresholds were adjusted to reflect realistic deal sizes and capital structures commonly seen in actual buyouts.

MidCap Filtering Criteria

Market Cap: \$750M to \$20B
→ Adjusted for mid-market deal sizes

EBITDA Margin ≥ 6.5%
→ Aligned with margins in mid-cap deals

ROA > 0
→ Retained to ensure baseline asset productivity

Debt/Equity < 5.5
→ Reflects typical leverage observed in mid-cap transactions

Sectors: Same 7 industries as Part I
→ Maintains focus on historically PE-active sectors

A total of **42** companies remained after applying all criteria

Legal & Structural Features (ML Inputs)

Regulatory Risk

Licensing requirements or oversight that may delay or prevent transaction approval

Contract Complexity

Key provisions in commercial agreements that directly affect control transfer and limit flexibility after closing

Debt Structure

Financing terms or covenants that constrain capital structure adjustments following acquisition

Execution Difficulty

Governance or shareholder structures that hinder execution or control transition

IP Dependence

Reliance on intangible assets with limited legal clarity or enforceability

These considerations reflect legal and structural constraints that can materially impact both deal execution and post-acquisition strategy

Labeling Strategy

Acquisition activity was labeled based on deal history between 2020 and 2025
The focus was on post-pandemic dynamics, where PE transaction patterns shifted in both volume and structure

Four tiers were assigned to reflect levels of acquisition progress

- 1.0 for completed acquisitions
- 0.7 for announced or formally engaged deals
- 0.3 for credible market speculation
- 0.0 for companies with no identifiable activity

This labeling enabled the model to prioritize targets along a continuous spectrum of buyout relevance

Legal and structural risks were also quantified and included in the model

Model Setup

A supervised learning model was trained to estimate the likelihood of acquisition

XGBoost Regressor was selected for its ability to capture non-linear patterns in structured financial data

The model used nine input features including four financial metrics and five legal or structural risk factors

Negative examples were selected from firms with weak profitability or excessive leverage

Model Results

The model generated acquisition scores for 42 companies that passed the screening process

Top-ranked firms showed strong profitability, efficient asset use and appropriate leverage

Model performance was stable, with mean squared error at **0.049** and R-squared at **0.66**

SHAP Interpretation

SHAP analysis shows how individual features contributed to a predicted score of 0.861

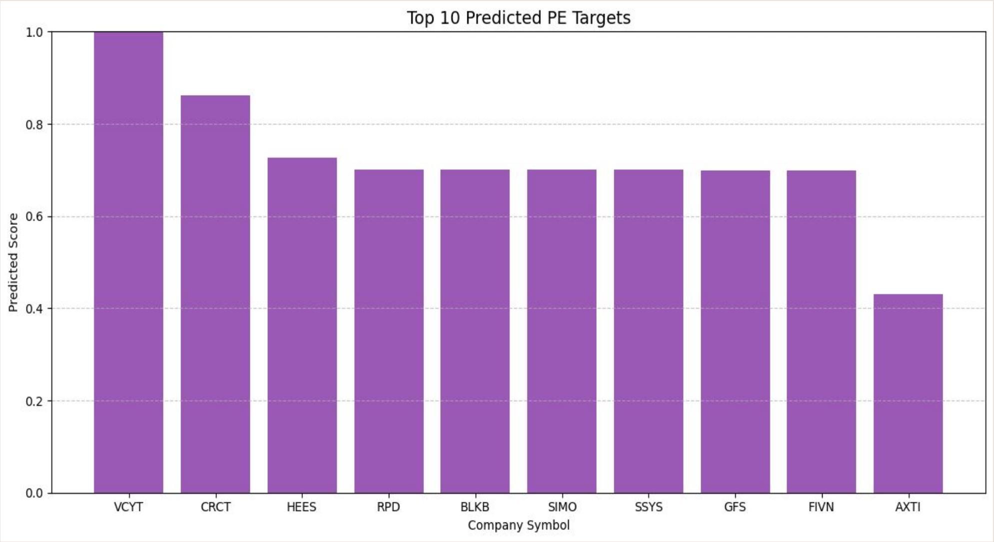
EBITDA margin and ROA were the most influential, followed by debt-to-equity and market cap

The plot explains how each input shifted the score from the model's baseline of 0.26

This interpretation reflects how the model captures key factors used in private equity assessment

Top Predicted PE Targets

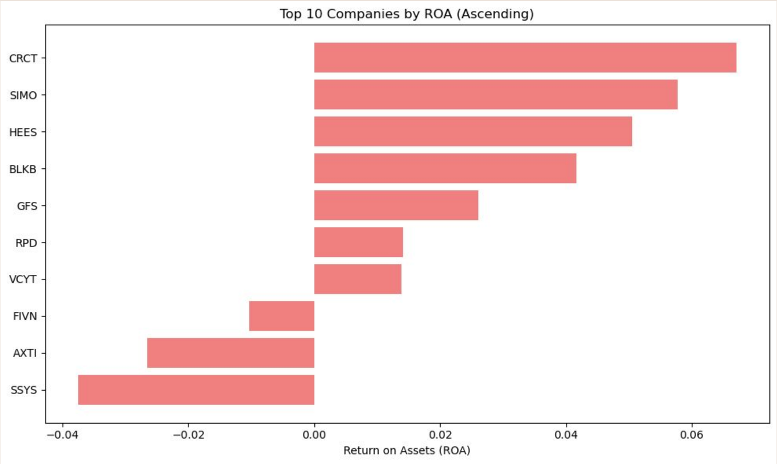
The model assigned acquisition scores to 42 screened companies
The top 10 firms ranked highest based on financial and structural fit
These candidates represent the most favorable profiles under the model's criteria

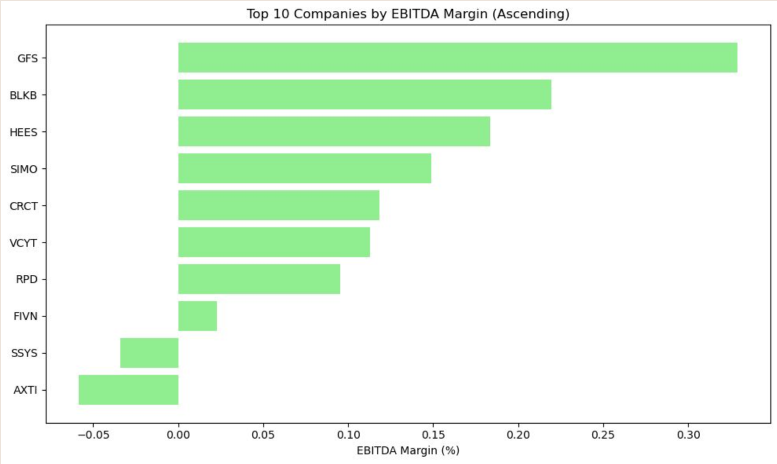


Financial Profiles of Top Candidates

The top-ranked companies also showed strong fundamentals in profitability and efficiency

The chart below compares return on assets and these patterns support the validity of the scoring logic





Final M&A Candidates

Among the top ten companies, three stand out based on model scores and market conditions

- Veracyte (VCYT)
 Diagnostics platform with strong potential for strategic acquisition in biotech and healthcare

- Rapid7 (RPD)
 Cybersecurity provider positioned for consolidation in enterprise cloud and security solutions

- Five9 (FIVN)
 Cloud-based contact center platform with prior acquisition interest and alignment with AI and automation trends

These candidates combine strong predictive scores with real-world acquisition momentum

Next Steps

The three shortlisted companies will move into focused review

Each will be assessed on strategic fit legal structure and deal feasibility

The process will support early stage screening and internal decision making

Thank you!