Portfolio Anomaly Detection

A Dual-Model Ensemble Approach for Investment Risk Assessment

Technical Whitepaper

December 2025

Tim Pinard
Software Engineer, Machine Learning

Version 1.0

Abstract

This paper presents a production-ready candidate machine learning system for detecting anomalous patterns in investment portfolios. The system employs a dual-model ensemble architecture combining an Autoencoder neural network with an Isolation Forest algorithm, each providing complementary detection capabilities. Trained on a universe of over 60 securities using 30 engineered technical features, the system identifies portfolios exhibiting unusual risk patterns that traditional metrics may miss. The Autoencoder achieves an F1 score of 0.63 with 93% recall on synthetic anomaly detection, demonstrating strong capability to catch true anomalies while maintaining reasonable precision. This work demonstrates practical viability of unsupervised learning techniques for financial risk assessment.

1. Introduction

1.1 The Challenge

Traditional portfolio risk metrics—Value at Risk (VaR), Sharpe ratio, and beta—measure *well-understood* risks based on historical patterns and statistical assumptions. However, these metrics often fail to detect emerging risks, unusual correlations, or portfolio compositions that deviate from historical norms.

This system addresses a different question: "*Given a portfolio of securities, is there something unusual about it that might indicate elevated risk?*" Rather than predicting specific outcomes, the system detects anomalies—portfolios that deviate from learned patterns of normal market behavior.

1.2 Key Innovations

- **Dual-Model Ensemble:** Combines unsupervised neural network pattern learning with statistical outlier detection
- **Universal Training:** Models trained on broad market universe can score any portfolio without portfolio-specific history
- **Feature-Rich Analysis:** 30 technical indicators capturing momentum, volatility, and market positioning
- **Market-Relative Scoring:** Z-score normalization contextualizes portfolio risk against current market conditions

2. System Architecture

2.1 Data Pipeline

The system processes market data through a structured pipeline: historical price and volume data for 60+ securities spanning 5 years is collected via public market data APIs (Yahoo Finance), then transformed into 30 technical indicators per security. These features are aggregated into portfolio-weighted vectors for model scoring.

2.2 The Two-Model Ensemble

The architecture employs two fundamentally different anomaly detection algorithms with complementary strengths:

Model	Type	Strength
Autoencoder	Neural Network	Complex non-linear pattern deviations
Isolation Forest	Tree-Based	Isolation of observations in sparse regions

The models are weighted 85% Autoencoder / 15% Isolation Forest based on validation performance. When both models agree on an anomaly, confidence is high.

3. Model Deep Dive: Autoencoder

3.1 Core Concept

The Autoencoder is a neural network trained to reconstruct its own input through a compressed bottleneck layer. This architecture forces the network to learn a lower-dimensional representation of normal market patterns. When presented with anomalous data, the reconstruction error spikes—the network cannot efficiently represent patterns it hasn't learned.

3.2 Architecture

Layer	Details
Input	30 features
Hidden Layer 1	128 neurons (ReLU + BatchNorm + Dropout 0.2)
Hidden Layer 2	64 neurons (ReLU + BatchNorm + Dropout 0.2)
Hidden Layer 3	32 neurons (ReLU + BatchNorm + Dropout 0.2)
Latent Space	**15 dimensions (50% compression)**
Decoder	Mirror of encoder: 32 → 64 → 128 → 30

The 50% compression ratio (30 features to 15 latent dimensions) forces meaningful representation learning. Batch normalization stabilizes training, while dropout (0.2) prevents memorization and forces redundant internal representations.

3.3 How this is meaningful

The 30 input features are highly correlated—the network discovers that it can represent momentum patterns with a single latent dimension capturing what's common across daily, weekly, and monthly returns. Anomalies, by definition, violate these learned correlations. If the network learns that high RSI typically accompanies high returns, a portfolio with high returns but low RSI produces a large reconstruction error on that feature.

4. Model Deep Dive: Isolation Forest

4.1 Core Concept

Isolation Forest operates on the basic principle: anomalies are easier to isolate than normal points. By randomly partitioning data with axis-aligned cuts, normal points (clustered in dense regions) require many cuts to isolate, while anomalies (in sparse regions) are isolated quickly. Shorter average path lengths correspond to higher anomaly scores.

4.2 Hyperparameters

Parameter	Value	Meaning
contamination	0.05	Expected proportion of anomalies in training data
n_estimators	200	Number of isolation trees in ensemble
max_samples	1024	Samples per tree (increased for better coverage)

4.3 Complementary Strengths

The two models have different failure modes, making them valuable as an ensemble:

Scenario	Autoencoder	Isolation Forest
Complex non-linear pattern shift	Catches it	May miss
Sharp statistical outlier	May miss if the pattern is learnable from the training data	Catches it
Gradual drift	Sensitive	Less sensitive
Noisy features	Can be confused	More robust

5. Feature Engineering

The system computes 30 technical indicators organized into seven categories, each capturing distinct aspects of market behavior:

5.1 Momentum (Returns)

Features: returns_1d, returns_5d, returns_20d, returns_60d

Multiple timeframes capture different market dynamics: daily noise, weekly momentum, monthly trends, and quarterly shifts. Divergence between timeframes is itself a signal—strong 5-day returns with weak 60-day returns suggests recent momentum fighting a larger trend.

5.2 Volatility

Features: volatility_20d, volatility_60d (annualized)

High volatility with low returns indicates market churn and indecision. Low volatility with high returns signals a strong trend. Values are annualized using the $\sqrt{252}$ factor.

5.3 Technical Indicators

RSI (rsi_14): Measures momentum exhaustion on a 0-100 scale. Values above 70 indicate overbought conditions; below 30 indicates oversold.

MACD (macd, macd_signal, macd_histogram): Three features capturing trend direction, confirmed direction, and acceleration. Divergence between MACD and price movement is a classic anomaly pattern.

Bollinger Bands (bb_upper, bb_middle, bb_lower, bb_position, bb_width): Position within volatility envelope and envelope width provide context for price movements relative to recent volatility.

5.4 Volume and Price Position

Volume: volume_ratio_20d, volume_ratio_60d, volume_std_20d, volume_std_60d

Price Position: price_to_52d_high, price_to_52d_low, price_to_200d_high,

price_to_200d_low, above_ma_50, above_ma_200

Volume patterns reveal market participation and conviction. Price position features provide trend context—a security near its 52-day high while above both moving averages indicates a strong uptrend.

6. Portfolio-Weighted Aggregation

A key architectural decision: individual security features are aggregated into a single portfolio-weighted feature vector before scoring. For each feature i:

$$portfolio\_feature[i] = \Sigma\ (weight\_j \times security\_j\_feature[i])$$

6.1 Design Rationale

Alternative Approach	Problem
Score each security, average scores	Misses portfolio-level interactions
Concatenate all features (30 × N)	Variable-length input; can't handle different sizes
Portfolio-weighted aggregation	**Fixed 30-dimensional representation; captures portfolio personality**

The portfolio-weighted approach captures portfolio "personality"—whether it's momentum-heavy, high-volatility, or fighting market trends. Anomalies occur when this personality is internally inconsistent or unusual compared to the training universe.

7. Model Validation

7.1 Validation Methodology

These scenarios are synthetically constructed but economically realistic, designed to reflect failure modes investors care about rather than purely statistical extremes:

- **Market Crash Portfolio:** Severe drawdowns, elevated volatility, oversold RSI
- **Momentum Bubble Portfolio:** Extreme positive returns, overbought RSI above 75
- **Volatility Spike Portfolio:** Sudden volatility expansion, abnormal trading volume
- **Correlation Breakdown Portfolio:** Diversification failure where positions move together
- **Liquidity Crisis Portfolio:** Volume collapse indicating exit difficulties

7.2 Performance Results

Metric	Autoencoder	Isolation Forest
F1 Score (Synthetic Anomalies)	**0.63**	0.19
Recall	**93%**	Lower
Model Agreement	**> 94%**	
Ensemble Weight	85%	15%

The Autoencoder's 93% recall indicates strong capability to catch true anomalies, critical for a risk detection system where false negatives have higher cost than false positives. Model agreement (>94%) indicates that, in the vast majority of cases, both models classify portfolios consistently as normal or anomalous.

8. Market-Relative Z-Score Analysis

Raw anomaly scores answer "Is my portfolio anomalous in absolute terms?" On volatile market days, everything appears somewhat anomalous. Z-score normalization provides market-relative, actionable context: "Is my portfolio more or less anomalous than the market today?"

For each evaluation, the portfolio is scored alongside the training universe. The z-score normalizes away market-wide stress:

$$z = (portfolio\_score - market\_mean) / market\_std$$

Z-Score	Interpretation	Action
≈ 0	Behaves like market average	Risk in line with market
+1 to +2	More anomalous than market	Elevated risk—monitor
> +2	**Much more anomalous**	**High excess risk—investigate**
-1 to -2	Less anomalous than market	Calmer than typical

9. Risk Classification Logic

The system combines model outputs with concentration metrics to produce actionable risk levels:

- **Critical:** Both models flag anomaly, OR Autoencoder anomaly with high concentration
- **High:** Autoencoder flags anomaly, OR ensemble score > 1.5, OR high concentration alone
- **Medium:** Isolation Forest flags anomaly (weaker standalone signal), OR elevated ensemble score with medium concentration
- **Low:** Within normal parameters across all dimensions

Concentration risk is assessed via Herfindahl-Hirschman Index (HHI) and maximum position weight, elevating risk for portfolios where single positions exceed 30% weight or HHI exceeds 0.25.

10. Future Directions

The current implementation provides a solid foundation for several enhancements:

- **Further tuning:** Feature and time period tuning, incorporating history events
- **LLM-Powered Explanations:** Natural language interpretation of anomaly patterns and personalized recommendations
- **Real-Time Monitoring:** WebSocket streaming for continuous portfolio surveillance
- **Historical Backtesting:** Validation against known market events (flash crashes, bubbles)
- **Portfolio-Specific Models:** Personalized anomaly detection as user history accumulates

11. Conclusion

This system demonstrates practical application of unsupervised machine learning to financial risk assessment. The dual-model ensemble approach provides robust anomaly detection by combining neural network pattern learning with statistical outlier identification. With 93% recall on synthetic anomalies and 94%+ model agreement, the system reliably identifies

portfolios needing further investigation.

The architecture's key strength is universality—models trained on broad market data can immediately score any portfolio without requiring portfolio-specific history. This enables rapid deployment and immediate value for new portfolios.

For organizations or individuals seeking to enhance their risk management capabilities with ML-driven insights, this approach offers a near production-ready foundation that complements rather than replaces traditional risk metrics.

References

1. Liu, F. T., Ting, K. M., & Zhou, Z. H. (2008). Isolation Forest. ICDM.
2. Kingma, D. P., & Welling, M. (2013). Auto-Encoding Variational Bayes. ICLR.
3. Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly Detection: A Survey. ACM Computing Surveys.
4. de Prado, M. L. (2018). Advances in Financial Machine Learning. Wiley.
5. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

About the Author

Tim Pinard is a software engineer with a keen interest in both machine learning and financial applications. The full project is available on GitHub at timpinard/portfolio-anomaly-detection.

Contact: t.pinard@gmail.com | Linked In