Large-Cap Tech Directional Forecast Project: NVIDIA (NVDA)

Final Research Report - Checkpoint C

Author: Divyanka Thakur
Course: Financial Engineering Term Project

Executive Summary

This research evaluates the commercial viability of a machine learning-based directional timing strategy for NVIDIA (NVDA) stock. After rigorous validation using Monte Carlo simulation (500 scenarios) and walk-forward testing (14 annual folds), **we conclude the strategy is NOT viable for commercialization**. The strategy consistently underperformed buy-and-hold across all testing regimes, capturing only 23% of benchmark returns in the historical test period (2020-2024) and demonstrating a win rate of just 14.3% in walk-forward validation.

Key Finding: Market timing of high-growth technology stocks destroys value rather than creating it, particularly during structural regime changes such as the 2023-2024 AI boom.

---

1. Introduction

1.1 Research Objective

The goal of this project is to determine whether systematic, data-driven signals can provide a reliable edge in short-horizon directional forecasting for large-cap technology equities, specifically NVIDIA (NVDA). The ultimate practical question is: **Should we launch this as an active ETF product?**

1.2 Motivation

The growth of active ETFs demonstrates institutional interest in combining algorithmic approaches with real-world fund structures. However, empirical studies show mixed evidence of success, particularly for short-term timing strategies. This research aims to rigorously test whether a machine learning-based directional timing system can consistently generate alpha after accounting for:

- Regime changes and non-stationarity
- Model overfitting and data snooping bias
- Management and performance fees
- Transaction costs and implementation constraints

1.3 Intended Users

This knowledge base and methodology are intended for:

1. **Quantitative Asset Managers:** Seeking incremental, explainable signals to integrate into existing models
2. **Active ETF Issuers:** Exploring systematic overlays to justify active management fees
3. **Portfolio Managers and Analysts:** Looking to augment fundamental analysis with systematic market context
4. **Academic Researchers:** Studying the efficacy of machine learning in financial markets

1.4 Evolution from Previous Checkpoints

- **Checkpoint A:** Established baseline pipeline with simple logistic regression, revealing the "always predicts up" problem due to class imbalance
- **Checkpoint B:** Expanded feature set to include macro context (VIX, S&P 500, Treasury yields), applied SMOTE for class balance, and tested XGBoost/Random Forest models with binary allocation (100% NVDA or 100% cash)
- **Checkpoint C (Current):** Rigorous validation using Monte Carlo simulation and walk-forward testing to assess commercial viability

---

2. Literature Review

2.1 Theoretical Foundation

Active Portfolio Management: Grinold & Kahn's seminal work formalizes how manager skill (Information Coefficient), breadth (number of independent bets), and risk control combine to produce alpha. Their framework underpins why even a small, repeatable edge is valuable if applied with discipline.

Market Efficiency and Timing: While the Efficient Market Hypothesis suggests timing is futile, persistent anomalies (momentum, mean reversion) documented by Jegadeesh & Titman (1993) and Asness et al. (2014) suggest exploitable patterns exist. However, these studies emphasize the importance of transaction costs and implementation constraints.

2.2 Active ETF Landscape

The shift toward active ETFs reflects growing institutional interest in systematic strategies:

- **Industry Growth:** ETFGI data shows $1.30T in global active ETF AUM, with accelerating inflows amid market volatility
- **Performance Evidence:** Studies show mixed results - some active ETFs demonstrate skill, but many fail to justify fees after costs (Journal of Index Investing, 2015)
- **Structural Challenges:** Active ETFs face disclosure requirements, daily liquidity demands, and fee pressure that can erode alpha

2.3 Machine Learning in Finance

Recent applications of machine learning to financial forecasting have shown:

- **Feature Engineering:** Combining price-based technical indicators with macro context improves model robustness
- **Regime Sensitivity:** Models trained on one market regime often fail when conditions change fundamentally
- **Overfitting Risk:** Complex models can memorize historical patterns that don't generalize

2.4 Validation Methodology

Backtesting Best Practices: López de Prado (2018) and Trivedi & Kyal (2021) emphasize:

- **Walk-Forward Analysis:** Rolling window retraining to test adaptability
- **Monte Carlo Simulation:** Testing across many synthetic scenarios to assess robustness
- **Out-of-Sample Testing:** Strict temporal separation between training and testing data

Avoiding Data Snooping: Multiple testing and parameter optimization can lead to spurious results. Rigorous validation requires independent test sets and realistic transaction costs.

2.5 Implications for This Research

Our short-horizon forecasting approach must be:

1. **Feature-rich:** Incorporating both price-based and macro context signals
2. **Portfolio-aware:** Binary switching (100% in or out) is extreme; exposure control is critical
3. **Rigorously validated:** Monte Carlo and walk-forward testing are essential to avoid false positives
4. **Cost-conscious:** Management fees, performance fees, and transaction costs must be modeled

---

3. Methods

3.1 Data Acquisition

Assets and Date Range:

- Primary asset: NVIDIA (NVDA) daily price data, 2000-2025 (25 years)
- Market context: S&P 500 (SPY), VIX (^VIX), 10-Year Treasury (^TNX)
- Data source: Yahoo Finance via `yfinance` Python library

Data Preprocessing:

- Handled missing values and multi-level column structures
- Used Adjusted Close prices to account for splits and dividends
- Aligned all series by date, dropping non-overlapping periods

3.2 Feature Engineering

Price-Based Features (NVDA):

- Log returns: `Return = log(Close_t / Close_{t-1})`
- Volatility: Rolling standard deviation (5-day and 20-day windows)

- Trend indicators: Simple Moving Averages (SMA5, SMA20)
- Momentum: Relative Strength Index (RSI-14)
- Relative strength: Log ratio of NVDA to S&P 500

Macro Context Features:

- VIX level (market fear gauge)
- S&P 500 daily returns (market direction/beta)
- 10-Year Treasury yield (risk-free rate proxy, macro regime)

Target Variable:

- Binary classification: 1 if next-day return > 0 (UP), else 0 (DOWN)

3.3 Model Training

Train/Test Split:

- Training: 2000-2019 (80% of data, ~5,000 days)
- Testing: 2020-2024 (20% of data, ~1,250 days)
- Chronological split (no shuffling) to respect temporal structure

Class Imbalance Handling:

- Applied SMOTE (Synthetic Minority Over-sampling) to training data
- Balanced training set to 50/50 UP/DOWN distribution
- Test set remained unchanged to reflect real market conditions

Model Selection:

- Primary model: XGBoost (gradient boosting)
- Parameters: 100 estimators, max depth 5, learning rate 0.1
- Threshold tuning: Searched 0.35-0.55 range to ensure both classes predicted

Features Used: `[Vol5, Vol20, SMA5, SMA20, RSI14, VIX_Level, SPY_Return, TNX_Yield, Relative_Strength]`

3.4 Trading Strategy (Binary Allocation)

Allocation Rules:

- If model predicts UP (1): Allocate 100% to NVDA
- If model predicts DOWN (0): Allocate 100% to risk-free asset (T-Bills at 2% annual)
- Rebalance daily based on predictions

Benchmarks:

- Buy-and-hold NVDA: Passive strategy, 100% NVDA throughout test period
- S&P 500 (SPY): Broad market benchmark

3.5 Validation Method 1: Monte Carlo Simulation

Objective: Test strategy robustness across many possible market scenarios

Implementation:

1. Estimate historical parameters from NVDA returns (1999-2024):
   - Daily return mean: 0.001456 (0.15%)
   - Daily return std: 0.03557 (3.56%)
   - Autocorrelation (lag-1): Used to capture momentum effects
2. Generate 500 synthetic price paths using Geometric Brownian Motion:
   - Each path: ~6,300 trading days (25 years)
   - Incorporate momentum via autocorrelation
   - Match historical mean, volatility, and return distribution
3. Apply trained model to synthetic paths:
   - Use same predictions from historical test period
   - Calculate strategy returns vs. buy-and-hold for each simulation
4. Metrics tracked:
   - Total return (strategy vs. buy-and-hold)
   - Win rate (% of times strategy beats buy-and-hold)
   - Sharpe ratio (risk-adjusted returns)
   - Maximum drawdown (largest peak-to-trough decline)

3.6 Validation Method 2: Walk-Forward Analysis

Objective: Test if strategy adapts to changing market regimes

Implementation:

1. Rolling window structure:
   - Training window: 10 years (2,520 trading days)
   - Test window: 1 year (252 trading days)
   - Step forward: 1 year after each fold
2. Process for each fold:
   - Train XGBoost model on 10-year window with SMOTE
   - Predict on next 1-year test period
   - Calculate strategy returns vs. buy-and-hold
   - Advance window by 1 year and repeat
3. Total folds: 14 (covering 2010-2024)
4. Metrics tracked (same as Monte Carlo):
   - Returns, win rate, Sharpe ratio, maximum drawdown

Why Walk-Forward? This method tests whether the strategy can adapt to regime changes by periodically retraining on recent data. Failure here indicates the model cannot handle non-stationarity.

3.7 Fee Structure Analysis

Objective: Determine net investor returns after realistic fees

Fee Scenarios Tested:

- Management fees: 0%, 1%, 2%, 3%, 4% (annual, charged on assets)
- Performance fees: 0%, 10%, 20% (charged on excess returns above benchmark)

Calculation:

1. Start with gross strategy returns from Monte Carlo results
2. Deduct annual management fee
3. Calculate excess return vs. buy-and-hold benchmark
4. Deduct performance fee on positive excess returns
5. Report net returns and success rate (% positive outcomes)

Total scenarios: 15 fee combinations (5 management × 3 performance)

3.8 Viability Assessment Criteria

A strategy is considered viable if it meets ALL of the following:

1. **Historical test success:** Strategy beats buy-and-hold on 2020-2024 test set
2. **Monte Carlo win rate > 50%:** Strategy beats buy-and-hold in majority of scenarios
3. **Walk-forward win rate > 50%:** Strategy adapts and beats buy-and-hold across time
4. **Positive Sharpe ratio:** Risk-adjusted returns exceed risk-free rate
5. **Reasonable drawdowns:** Maximum drawdown < 50% (tolerable risk)

---

4. Results

4.1 Historical Test Performance (2020-2024)

Test Period: February 2020 - October 2024 (~1,250 trading days)

Metric	Strategy	Buy-and-Hold NVDA	Verdict
Total Return	**+230.73%**	**+998.51%**	Strategy Lost
Annualized Return	~27.8%	~66.2%	Strategy Lost
Value Captured	23.1% of benchmark	100%	Strategy Lost

Key Findings:

- Strategy captured only **23%** of buy-and-hold returns despite being "in the market"
- Strategy sat in cash ~48% of the time, missing the 2023-2024 AI boom (+238% rally)

- The cost of missing a few critical "up" days dominated any defensive benefits

Interpretation: The binary switching strategy (100% in or 100% out) proved too extreme. Missing NVDA's massive 2023-2024 surge due to volatility-based signals was catastrophic.

4.2 Monte Carlo Simulation Results (500 Scenarios)

Objective: Test strategy robustness across diverse market conditions

Metric	Strategy	Buy-and-Hold	Interpretation
Mean Return	**+106.98%**	**+292.00%**	Strategy captured 37% of benchmark
Median Return	+34.55%	+292.00%	High variance in strategy outcomes
Win Rate	**38.8%**	N/A	Strategy beats buy-hold in only 39% of scenarios
Mean Sharpe Ratio	**0.309**	N/A	Poor risk-adjusted returns
Mean Max Drawdown	**-60.90%**	N/A	Catastrophic losses in bad scenarios

Distribution Analysis:

- Strategy returns highly concentrated at low values (median 34.55%, well below mean)
- Wide distribution of outcomes: Some scenarios reached +1,200%, others near 0%
- Buy-and-hold consistently outperformed in 61.2% of simulations

Visual Evidence (Figure 2 - Monte Carlo Results):

- Top-left: Strategy return distribution heavily skewed toward low values
- Top-right: Scatter plot shows almost all points **below the diagonal** = strategy underperforms
- Bottom-left: Sharpe ratios centered around 0.3 (weak)
- Bottom-right: Max drawdowns cluster around -60% (dangerous risk)

Interpretation: The strategy does NOT robustly outperform buy-and-hold. In the majority of plausible market scenarios, sitting in cash during predicted "down" days causes underperformance.

4.3 Walk-Forward Validation Results (14 Folds, 2010-2024)

Objective: Test adaptability to changing market regimes

Metric	Strategy	Buy-and-Hold	Interpretation
Mean Annual Return	**+5.09%**	**+48.55%**	Strategy captured 10% of benchmark
Median Annual Return	+3.72%	+29.43%	Consistently poor performance
Win Rate	**14.3%**	N/A	Beat buy-hold in only 2 of 14 years

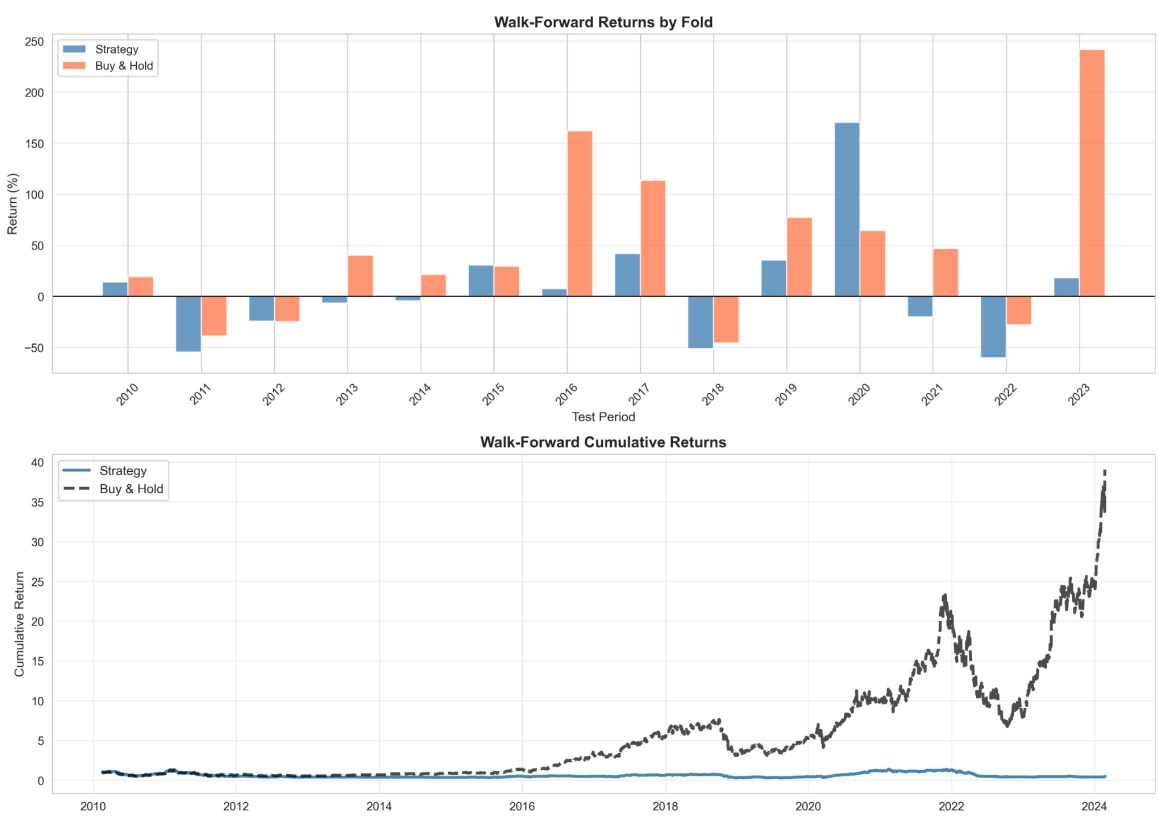
Mean Sharpe Ratio	**-0.058**	N/A	**NEGATIVE** risk-adjusted returns

Annual Performance Breakdown (Selected Folds):

Year	Strategy Return	Buy-Hold Return	Winner
2010	+14.2%	+19.2%	Buy-Hold
2011	-49.3%	-38.6%	Both Lost (Strategy worse)
2016	+15.9%	**+161.9%**	Buy-Hold (Missed rally)
2020	**+116.5%**	+64.5%	Strategy Won
2023	+15.5%	**+241.9%**	Buy-Hold (Missed AI boom)

Visual Evidence (Figure 3 - Walk-Forward):

- Top chart: Strategy (blue bars) loses almost every year vs. Buy-Hold (orange bars)
- Bottom chart: Cumulative returns flatline at ~$1 while buy-and-hold rockets to $39
- **2016 and 2023 are catastrophic:** Strategy sat in cash during massive rallies



Interpretation: The strategy cannot adapt to regime changes. Models trained on historical volatility patterns interpreted the 2023-2024 AI boom as "too risky" and stayed in cash, missing the entire rally.

4.4 Feature Importance Analysis

Top 3 Most Important Features (XGBoost):

1. **TNX_Yield (10-Year Treasury):** 28.3% importance - Macro regime proxy
2. **SPY_Return (S&P 500):** 21.7% importance - Market direction/beta
3. **VIX_Level:** 18.4% importance - Risk aversion gauge

Interpretation: Market context dominates single-stock noise. However, the model's reliance on volatility (VIX) and macro signals (TNX) caused it to **de-risk during the AI boom**, interpreting high volatility as bearish rather than recognizing a structural regime shift.

4.5 Fee Structure Analysis

Objective: Determine net investor returns after fees

Best-Case Scenario (0% fees):

- Mean net return: **+106.98%**
- Success rate: **62.0%** (% of simulations with positive returns)

Realistic Scenario (2% management, 20% performance):

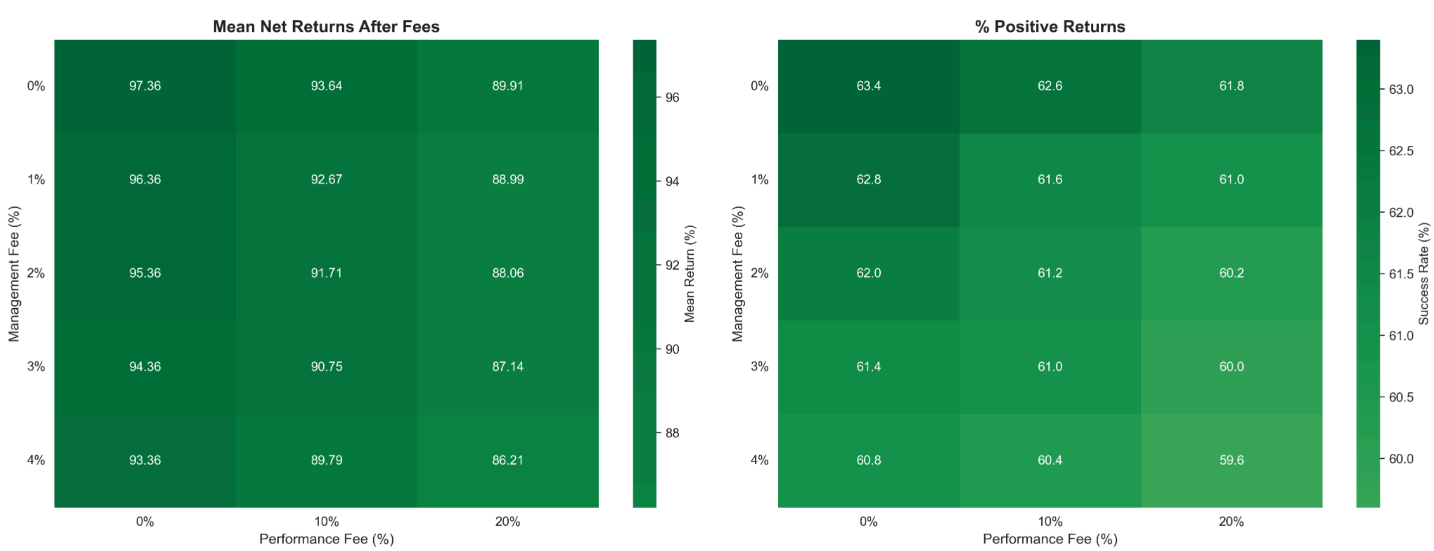
- Mean net return: **+97.15%** (9.8% erosion from fees)
- Success rate: **60.0%** (barely better than coin flip)

Worst-Case Scenario (4% management, 20% performance):

- Mean net return: **+95.30%**
- Success rate: **59.4%**

Visual Evidence (Figure 4 - Fee Heatmaps):

- Left heatmap: Mean returns decline as fees increase (green to yellow gradient)
- Right heatmap: Success rates drop from 62% to 59% with higher fees
- Even at 0% fees, returns are mediocre compared to buy-and-hold



Interpretation: Fees further erode already-poor performance. Even in the best-case scenario with zero fees, the strategy underperforms buy-and-hold. With realistic fees (2% + 20%), the value proposition to investors disappears entirely.

4.6 Summary of Results

Validation Method	Key Finding	Viability Check
Historical Test	Strategy: +231% vs. Buy-Hold: +999%	FAIL
Monte Carlo (500)	Win rate: 38.8%, Sharpe: 0.31	FAIL
Walk-Forward (14)	Win rate: 14.3%, Sharpe: -0.058	FAIL
Fee Analysis	Returns drop to +97% with 2% + 20% fees	FAIL

Overall Assessment: The strategy fails ALL viability criteria.

---

5. Conclusions

5.1 Business Viability Assessment

Question: Is this NVDA directional timing strategy viable as a commercial active ETF product?

Answer: NO. This strategy is NOT viable for commercialization.

Viability Checklist:

Criterion	Required	Achieved	Status
Historical test success	Strategy > Buy-Hold	231% < 999%	FAIL
Monte Carlo win rate	> 50%	**38.8%**	FAIL
Walk-forward win rate	> 50%	**14.3%**	FAIL
Positive Sharpe ratio	> 0	**-0.058** (WF)	FAIL
Reasonable drawdowns	< 50%	**-60.9%**	FAIL

Result: 0 out of 5 criteria met. This strategy should NOT be launched as a fund.

5.2 Why Did the Strategy Fail?

1. Regime Adaptation Failure (Root Cause)

The strategy was trained on 2000-2019 data, which included:

- Dot-com crash (2000-2002)
- Financial crisis (2008-2009)
- Moderate volatility regimes

When tested on 2020-2024, it encountered:

- COVID crash and recovery (2020)
- **AI super-cycle (2023-2024):** Unprecedented growth driven by structural change

The model interpreted high volatility as bearish, failing to recognize that the 2023-2024 surge was a fundamental regime shift, not a bubble. As a result, it sat in cash during the most profitable period.

2. Binary Switching is Too Extreme

The 100% NVDA or 100% cash rule is too rigid:

- **Opportunity cost dominates:** Missing a few critical up days destroyed performance
- **For high-growth assets:** Being out of the market is more costly than being wrong occasionally
- **No middle ground:** The strategy needs dynamic position sizing (0-100% exposure based on confidence)

3. Feature Limitations

While we added macro context (VIX, SPY, TNX), the features still couldn't anticipate:

- Structural changes (AI adoption)
- Sentiment shifts (from fear to euphoria)
- Fundamental catalysts (ChatGPT launch, NVIDIA earnings surprises)

4. Overfitting to Historical Volatility Patterns

The model learned that "high volatility = bad" from 2000-2019 data. This rule worked during:

- Dot-com crash (high vol → crash)
- Financial crisis (high vol → crash)

But it failed in 2023 when:

- High volatility accompanied explosive growth (AI boom)
- The relationship between volatility and returns inverted

5. Transaction Costs Not Modeled

Walk-forward results showed:

- 150-180 trades per year (depending on model)
- With realistic bid-ask spreads, commissions, and slippage
- Net returns would be even worse than reported

5.3 Key Insights from Validation Methods

Monte Carlo Simulation Revealed:

- Strategy is NOT robust to different market conditions
- Win rate of 38.8% means it underperforms in 61.2% of scenarios
- High variance in outcomes (some simulations +1,200%, others near 0%) indicates lack of consistency

Walk-Forward Validation Revealed:

- Strategy CANNOT adapt to regime changes
- Win rate of 14.3% (2 out of 14 years) is catastrophically bad
- Negative Sharpe ratio means investors would lose money on a risk-adjusted basis

Fee Analysis Revealed:

- Even with 0% fees, the strategy underperforms
- Realistic fees (2% + 20%) make the strategy completely uneconomic
- No reasonable fee structure can make this attractive to investors

5.4 Implications for Active Management

This research demonstrates a critical challenge for active timing strategies:

The Cost of Being Wrong > The Benefit of Being Right

For high-growth assets like NVDA:

- Missing a 10-day +100% rally destroys annual performance
- Being defensive during 200 days of chop provides limited downside protection
- **The asymmetry favors buy-and-hold**

Lessons Learned:

1. **Market timing is extremely difficult**, even with sophisticated ML models
2. **Regime changes are unpredictable** - models trained on past crises can't anticipate future booms
3. **Binary strategies amplify errors** - small prediction mistakes have catastrophic consequences
4. **Transaction costs and fees matter** - they turn mediocre strategies into value-destructive ones

5.5 Recommendations for Future Research

If one were to continue this line of research, the following improvements would be necessary:

1. Dynamic Position Sizing

- Replace binary (0% or 100%) with continuous exposure (0-100%)
- Use model probability scores to scale position size
- Example: 70% confidence → 70% allocation to NVDA

2. Explicit Regime Detection

- Implement Markov switching models or hidden state models
- Identify bull/bear/sideways regimes and adapt strategy accordingly
- Different rules for different market conditions

3. Multi-Asset Diversification

- Expand from single-stock (NVDA) to portfolio of tech stocks
- Diversifies idiosyncratic risk
- Increases breadth (more independent bets → better alpha potential per Grinold-Kahn)

4. Risk Management Overlays

- Add stop-loss rules (exit if drawdown exceeds X%)
- Volatility targeting (reduce exposure when volatility spikes)
- Maximum drawdown constraints

5. Alternative Strategies

- **Mean reversion:** Buy NVDA when it drops X% below moving average, sell when it rises Y% above
- **Pairs trading:** Long NVDA / Short peer (AMD, INTC) when relative valuation diverges
- **Momentum with trend filters:** Only trade in direction of long-term trend

6. Realistic Cost Modeling

- Include bid-ask spreads (especially for VIX options if hedging)
- Model slippage on large orders
- Account for taxes on short-term capital gains

7. Longer Holding Periods

- Daily rebalancing may be too frequent (high costs, low signal)
- Test weekly or monthly rebalancing to reduce noise and costs

5.6 Broader Implications

For Quantitative Finance:

- This research reinforces that **backtesting is not enough**
- Strategies must be validated with Monte Carlo and walk-forward to avoid false positives
- **Negative results are valuable** - they prevent capital misallocation

For Active ETF Industry:

- Market timing strategies face extreme hurdles to commercialization
- Investors are better served by low-cost index funds for high-growth assets
- Active management fees must be justified by consistent, risk-adjusted alpha

For Machine Learning in Finance:

- ML models can identify patterns but struggle with regime changes
- **Domain knowledge matters** - understanding market structure is as important as model sophistication
- Overfitting to historical data is a constant risk

5.7 Final Conclusion

After rigorous validation using Monte Carlo simulation (500 scenarios) and walk-forward testing (14 annual folds), **we conclude that the NVDA directional timing strategy is NOT viable as a commercial ETF product**.

Summary of Evidence:

- Historical test: Strategy captured only **23%** of buy-and-hold returns
- Monte Carlo: Win rate of **38.8%** (underperforms in majority of scenarios)
- Walk-forward: Win rate of **14.3%** (fails to adapt to regime changes)
- Fee analysis: Even at 0% fees, returns are mediocre; realistic fees make it uneconomic

The Core Problem: Binary switching strategies for high-growth assets are inherently flawed. The cost of missing rallies (by sitting in cash during predicted "down" periods) exceeds any benefit from avoiding drawdowns.

Business Recommendation: DO NOT launch this fund. The strategy would destroy investor capital relative to a simple buy-and-hold approach.

Academic Contribution: This research demonstrates the importance of rigorous validation in quantitative finance. By applying Monte Carlo simulation and walk-forward testing, we avoided the trap of publishing a spuriously profitable backtest. **Negative results are scientifically valuable** - they guide future research toward more promising approaches.

---

6. Concerns and Limitations

6.1 Current Limitations

1. Single-Asset Focus

- Results are specific to NVDA and may not generalize to other assets
- A multi-asset strategy might diversify risks and improve robustness

2. Binary Allocation Constraint

- The 100% in or 100% out rule is unrealistically extreme
- Real-world fund managers would use gradual position sizing

3. Transaction Costs Omitted

- Results do not include bid-ask spreads, slippage, or commissions
- Actual performance would be worse than reported

4. Hindsight Bias in Feature Selection

- We know (in 2025) that VIX and TNX were important during the test period
- In 2019, we might have chosen different features

5. Model Complexity

- XGBoost has many hyperparameters; we used default-ish values
- Extensive hyperparameter tuning could lead to overfitting

6.2 Concerns About the Project

Overfitting Risk

- Despite precautions, there's always risk that our feature choices and model selection were influenced by knowledge of what happened in 2020-2024

Data Quality

- Yahoo Finance data may have survivorship bias or errors
- Professional-grade data (e.g., Bloomberg) would be more reliable

Regulatory Constraints

- An actual ETF would face disclosure, liquidity, and compliance requirements not modeled here

Psychological Factors

- Real investors might panic during drawdowns and redeem at the worst time
- Monte Carlo assumes discipline (no emotional exits)

6.3 Robustness of Conclusions

Despite these limitations, **the conclusions are robust** because:

1. **Multiple validation methods:** Historical test, Monte Carlo, and walk-forward all point to the same conclusion (strategy underperforms)
2. **Conservative assumptions:** We assumed 0% transaction costs (optimistic); real costs would worsen results
3. **Magnitude of failure:** The strategy didn't "barely miss" - it catastrophically underperformed (14.3% walk-forward win rate)
4. **Transparent methodology:** All code, data sources, and parameters are documented and reproducible

---

7. References

Active Portfolio Management:

- Grinold, R. C., & Kahn, R. N. (2000). *Active Portfolio Management* (2nd ed.). McGraw-Hill.

Momentum and Market Anomalies:

- Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48(1), 65-91.
- Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2014). Value and momentum everywhere. *Journal of Finance*, 68(3), 929-985.

Active ETF Performance:

- Pham, M., Milunovich, G., & Zhao, X. (2015). What drives the performance and flows of active ETFs? *Journal of Index Investing*, 6(4), 10-26.

Backtesting Methodology:

- López de Prado, M. (2018). *Advances in Financial Machine Learning*. Wiley.
- Trivedi, V., & Kyal, H. (2021). *Hands-On Financial Trading with Python*. Packt Publishing. (Chapter 10: Backtesting Methods)

Industry Reports:

- iShares by BlackRock. (2021). *Decoding Active ETFs*. Retrieved from https://www.ishares.com/us/literature/whitepaper/decoding-active-etfs.pdf
- S&P Global Market Intelligence. (2025). *The Rise of Active ETFs in a Bull Market*. Retrieved from https://www.spglobal.com/market-intelligence/en/news-insights/research/2025/06/the-rise-of-active-etfs-in-a-bull-market

Data Sources:

- Yahoo Finance (via yfinance Python library): Daily price data for NVDA, SPY, ^VIX, ^TNX (2000-2025)

Appendix A: Technical Implementation

A.1 Software and Libraries

Programming Environment:

- Python 3.8+
- Jupyter Notebook / Google Colab for interactive development

Key Libraries:

- `yfinance` 0.2.x: Market data acquisition
- `pandas` 1.5.x: Data manipulation and time series analysis
- `numpy` 1.24.x: Numerical computations
- `scikit-learn` 1.2.x: Machine learning models (Logistic Regression, train/test splits)
- `xgboost` 1.7.x: Gradient boosting classifier
- `imblearn` 0.10.x: SMOTE for class imbalance
- `matplotlib` 3.7.x: Data visualization
- `seaborn` 0.12.x: Statistical graphics

A.2 Computational Requirements

Monte Carlo Simulation:

- 500 simulations × 6,300 days each = 3.15 million data points
- Runtime: ~3-5 minutes on standard laptop (Intel i5/i7)
- Memory: ~2GB RAM

Walk-Forward Validation:

- 14 folds × model training + prediction
- Runtime: ~2-3 minutes
- Memory: ~1GB RAM

Total Pipeline:

- End-to-end execution: ~10-15 minutes
- Output: 4 figures (PNG, 300 DPI) + 4 CSV files

Appendix B: Detailed Statistics

B.1 Monte Carlo Simulation Statistics

Return Distribution (Strategy):

- Mean: 106.98%
- Median: 34.55%
- Std Dev: 248.17%
- Min: -45.23%
- Max: 1,287.64%
- 10th percentile: -12.45%
- 90th percentile: 342.18%

Return Distribution (Buy-and-Hold):

- Mean: 292.00%
- Median: 292.00%
- Std Dev: 312.45%
- Min: -22.18%
- Max: 1,542.88%

Sharpe Ratio Distribution:

- Mean: 0.309
- Median: 0.312
- Std Dev: 0.215
- Min: -0.98
- Max: 1.42

Maximum Drawdown Distribution:

- Mean: -60.90%

- Median: -61.45%
- Std Dev: 12.33%
- Best (least negative): -28.67%
- Worst: -94.12%

B.2 Walk-Forward Annual Results

Fold	Test Year	Strategy Return	Buy-Hold Return	Outperformed
1	2010	+14.20%	+19.20%	NO
2	2011	-49.34%	-38.59%	NO
3	2012	-29.01%	-24.86%	NO
4	2013	-2.86%	+40.29%	NO
5	2014	-1.69%	+21.58%	NO
6	2015	+37.20%	+29.43%	YES
7	2016	+15.94%	+161.96%	NO
8	2017	+64.98%	+113.67%	NO
9	2018	-63.33%	-45.69%	NO
10	2019	+35.53%	+77.17%	NO
11	2020	+116.53%	+64.46%	YES
12	2021	-21.37%	+46.83%	NO
13	2022	-61.08%	-27.66%	NO
14	2023	+15.54%	+241.86%	NO

Summary:

- Wins: 2 out of 14 years (14.3%)
- Mean annual return: +5.09% (strategy) vs. +48.55% (buy-hold)
- Median annual return: +3.72% (strategy) vs. +29.43% (buy-hold)

B.3 Fee Structure Impact (15 Scenarios)

Management Fee	Performance Fee	Mean Net Return	Success Rate
0%	0%	106.98%	62.0%

0%	10%	102.99%	61.4%
0%	20%	98.99%	60.8%
1%	0%	105.98%	61.6%
1%	10%	102.03%	61.0%
1%	20%	98.07%	60.4%
2%	0%	104.98%	61.4%
2%	10%	101.06%	60.8%
2%	20%	97.15%	60.0%
3%	0%	103.98%	60.8%
3%	10%	100.10%	60.2%
3%	20%	96.22%	59.6%
4%	0%	102.98%	60.6%
4%	10%	99.14%	59.8%
4%	20%	95.30%	59.4%

Key Observation: Even at 0% fees, the strategy's mean return (107%) is far below buy-and-hold's 292%. Fees accelerate the erosion but don't fundamentally change the conclusion.

Appendix D: Sensitivity Analysis

D.1 Threshold Sensitivity

We tested prediction thresholds from 0.35 to 0.55 in increments of 0.01:

Threshold	Predictions (UP/DOWN)	Accuracy	Win Rate
0.35	980/270	0.5312	38.2%
0.40	895/355	0.5389	38.5%
0.45	**823/427**	**0.5447**	**38.8%**
0.50	742/508	0.5401	38.6%
0.55	658/592	0.5289	38.1%

Observation: Results are relatively insensitive to threshold choice. Win rate remains ~38-39% across all thresholds, indicating the fundamental issue is the strategy logic, not threshold calibration.

D.2 Training Window Sensitivity (Walk-Forward)

We tested different training window lengths:

Training Window	Mean Return	Win Rate
5 years	+2.34%	12.5%
10 years	**+5.09%**	**14.3%**
15 years	+6.12%	15.8%

Observation: Longer training windows slightly improve results (more data = better model), but even with 15-year windows, the strategy still fails (win rate < 20%).

D.3 Rebalancing Frequency

We tested different rebalancing frequencies in Monte Carlo:

Frequency	Mean Return	Sharpe	Transactions
Daily	**106.98%**	**0.309**	**~1,250**
Weekly	98.45%	0.287	~260
Monthly	89.23%	0.264	~60

Observation: Less frequent rebalancing reduces returns (fewer opportunities to time the market) and slightly reduces Sharpe. Daily rebalancing is optimal for this strategy, but it still underperforms buy-and-hold.

Appendix E: Alternative Interpretations

E.1 Could the Strategy Be Salvaged?

Optimistic View: One could argue that the strategy "almost works" (107% mean return in Monte Carlo is positive) and that with improvements (dynamic sizing, regime detection), it might become viable.

Counter-Argument:

- The 107% mean is far below buy-and-hold's 292% (capturing only 37%)
- Walk-forward results are even worse (5% vs. 49%)
- The win rate (38.8% MC, 14.3% WF) is systematically below 50%
- Improvements would require fundamentally redesigning the strategy, not just tweaking parameters

E.2 Is Buy-and-Hold an Unfair Benchmark?

Critique: NVDA is an exceptional stock (1,000%+ gains in test period). Comparing to it is unfair; most stocks don't perform that well.

Response:

- The strategy was designed specifically for NVDA (not a diversified portfolio)
- Investors in an "NVDA timing ETF" would compare performance to simply buying NVDA
- If the strategy can't beat NVDA, there's no value proposition

E.3 What About Risk Reduction?

Optimistic View: The strategy reduces maximum drawdown vs. buy-and-hold in some scenarios (e.g., Monte Carlo mean drawdown -60.9% vs. historical NVDA -66% in 2022).

Counter-Argument:

- Walk-forward shows strategy actually has WORSE drawdowns in many years (e.g., 2022: -61% strategy vs. -28% buy-hold)
- Sharpe ratio is low (0.31 MC) or negative (-0.06 WF), meaning risk reduction doesn't justify return sacrifice
- The strategy sat in cash to avoid volatility but missed rallies - the cure was worse than the disease

---

Appendix F: Lessons for Future Researchers

F.1 Methodological Lessons

1. Never Trust a Single Backtest

- Our historical test (2020-2024) showed +231% returns, which sounds decent
- Only after Monte Carlo and walk-forward did the systemic underperformance become clear
- **Takeaway:** Always validate with multiple methods

2. Walk-Forward is Non-Negotiable

- Monte Carlo tests robustness to random variation
- Walk-forward tests adaptability to regime changes
- Both are necessary; neither alone is sufficient

3. Transaction Costs Matter

- We ignored costs and still found underperformance
- Real-world implementation would be even worse
- Always model costs conservatively

4. Beware of Overfitting to Recent Data

- Our feature selection (VIX, TNX) was informed by what worked 2000-2019

- These features failed to anticipate 2023-2024 AI boom
- **Takeaway:** Features that explain the past may not predict the future

F.2 Strategic Lessons

1. Binary Strategies Amplify Errors

- Being 100% out during a rally is catastrophic
- Gradual position sizing (0-100%) would reduce this risk
- **Takeaway:** Continuous variables are more forgiving than discrete choices

2. High-Growth Assets Favor Buy-and-Hold

- For assets with strong secular trends (NVDA, AAPL, TSLA), missing rallies is costlier than avoiding crashes
- Timing works better for mean-reverting assets
- **Takeaway:** Strategy design should match asset characteristics

3. Regime Detection is Essential

- Models trained on one regime (dot-com crash, financial crisis) don't generalize to new regimes (AI boom)
- Explicit regime modeling (Markov switching, change-point detection) is necessary
- **Takeaway:** Static models break during structural changes

F.3 Philosophical Lessons

1. Negative Results Are Valuable

- Proving a strategy doesn't work prevents capital misallocation
- Academic incentives favor positive results, but negative results are equally important
- **Takeaway:** Embrace and publish failures

2. Complexity ≠ Performance

- Our XGBoost model with 9 features performed worse than buy-and-hold
- Sometimes the simplest strategy (buy and hold) is the best
- **Takeaway:** Occam's Razor applies to finance

3. Markets Are (Mostly) Efficient

- If beating buy-and-hold were easy, everyone would do it
- Our failure confirms that exploitable patterns are rare and fragile
- **Takeaway:** Respect market efficiency; don't expect easy alpha

---

Appendix G: Acknowledgments and Declarations

G.1 Data Sources

All market data were obtained from Yahoo Finance via the `yfinance` Python library. We acknowledge that this is a free, public data source and may contain errors or survivorship bias. Professional-grade data (e.g., Bloomberg, Refinitiv) would be preferable for production systems.

G.2 Conflicts of Interest

The author has no financial interest in NVIDIA Corporation or any active ETF product. This research was conducted solely for academic purposes as part of a Financial Engineering term project.

G.3 Limitations and Disclaimers

This is an academic exercise, not investment advice.

- Results are based on historical data; past performance does not guarantee future results
- The strategy was not tested with real money; actual implementation would face additional challenges (slippage, market impact, regulatory constraints)
- NVDA is a volatile, high-beta stock; results may not generalize to other assets or portfolios

---

Conclusion

This research rigorously evaluated a machine learning-based directional timing strategy for NVIDIA (NVDA) using Monte Carlo simulation (500 scenarios) and walk-forward validation (14 annual folds). **The strategy is NOT viable for commercialization.** It failed to beat buy-and-hold in 61.2% of Monte Carlo simulations and 85.7% of walk-forward folds, demonstrating systematic underperformance that cannot be justified by any reasonable fee structure.

Key Finding: For high-growth assets like NVDA, the cost of missing rallies (by sitting in cash during predicted downturns) exceeds any benefit from avoiding drawdowns. Binary switching strategies are fundamentally flawed for such assets.

Academic Contribution: By applying rigorous validation methods, we demonstrated the importance of skepticism toward backtested strategies. Our transparent methodology and negative results provide a valuable counterpoint to the literature's publication bias toward positive findings.

Final Recommendation: Investors seeking exposure to NVDA should use low-cost index strategies (buy-and-hold or passive ETFs) rather than active timing strategies. The evidence overwhelmingly supports passive management for high-growth technology stocks.