Decision Trees for Market Risk Detection: Predicting Drawdowns and Portfolio Protection

1st Samuel Salama
School of Informatics, Computing, and Engineering
Indiana University
Bloomington, USA
ssalama@iu.edu

***Abstract*—This project develops an iterative decision tree-based system for detecting market risks in 3 major S&P 500 stocks (AAPL, MSFT, GOOGL), evolving from a basic price prediction model to an accurate risk detection framework. Through five major iterations, I explored various feature combinations and model designs, ultimately achieving 99% accuracy in risk detection. The system employs engineered features focusing on drawdown characteristics, volatility patterns, and price momentum indicators. The final model achieved balanced precision-recall metrics (81% for both) in identifying risk events, demonstrating significant improvement over traditional technical indicator-based approaches. This research provides a potential framework for portfolio protection while maintaining low false-positive rates, making it suitable for real-world applications in risk management.**

***Index Terms*—decision trees, risk detection, market drawdown, portfolio protection, machine learning, algorithmic trading, S&P 500**

I. Introduction

Financial market risk management is a challenge that is constantly increasing in complexity and importance, and while research has explored market prediction models, there always remains room for improvement in developing reliable approaches to risk detection and market drawdown prediction. Recent systematic reviews have highlighted that "decision making" is the most critical step repeated throughout the risk management process (Zaku & Uysal, 2022), yet traditional approaches often fail to adequately integrate decision-making under uncertainty. While risk assessment and management was established as a scientific field 30-40 years ago (Aven, 2016), the application of decision trees for systematic risk detection represents a relatively new and promising direction.

My research addresses these three challenges that were identified in recent literature:

- Decision trees must be integrated with other methods to effectively handle uncertainty in risk management (Zaku & Uysal, 2022). Recent systematic reviews show that over 81% of relevant research in this area has been published in the last five years, indicating growing recognition of this need.
- The financial domain requires quantification of uncertainty in risk management processes (Aven, 2016). Decision trees can help provide a structured approach for this quantification, specifically when combined with other machine learning techniques.
- Risk detection needs to span identification, analysis, and monitoring phases (Zaku & Uysal, 2022). However, current literature still shows some gaps in risk monitoring and control processes compared to other risk management phases.

To address some of these challenges my approaches make several contributions:

- Development of an iterative decision tree framework that evolves from basic market prediction (38% accuracy) to risk detection (99% accuracy), demonstrating how traditional technical analysis can be transformed into effective risk management tools.
- Feature engineering that progressed from basic technical indicators to focused risk metrics, leading to a balanced set of indicators (Drawdown: 58.3%, Drawdown_Speed: 25.2%, Volatility_Ratio: 16.5%) specifically designed for risk detection.
- Framework for market risk detection that achieves balanced precision and recall (both 81%), addressing the critical need for reliable early warning systems in financial markets.

This work builds upon recent advances and ideas in risk assessment methodologies while addressing some limitations identified in current approaches. My framework incorporates findings from systematic reviews of decision tree applications in risk management (Zaku & Uysal, 2022), particularly in the areas of feature selection, model validation, and risk characterization.

The remainder of this paper is organized as follows: Section II presents my methodology and theoretical framework, Section III describes the iterative development of my risk detection model, Section IV presents experimental results and validation, and Section V discusses implications and future directions. My approach demonstrates how systematic risk assessment principles can be effectively applied to market risk detection, providing both theoretical insights and practical tools for risk management.

II. Methods

My methodology consists of data collection, feature engineering, model development, and validation processes designed to help create an accurate risk detection framework.

My approach evolved through multiple iterations, each refining the model's ability to identify market risks while maintaining applicability.

A. Data Collection and Preprocessing

My program utilized historical market data from three major S&P 500 stocks (AAPL, MSFT, GOOGL) spanning from January 2019 to December 2023. Data was sourced through the Yahoo Finance API and then exported to its own csv file. The dataset consisted of daily price data including Open, High, Low, Close, Adjusted Close, and Volume metrics. The initial dataset comprised 3,714 trading days across all stocks, with no missing values, ensuring data completeness and consistency.

B. Feature Engineering

The feature engineering process progressed through several stages, but ultimately focused on three primary categories of risk indicators:

1) Drawdown Metrics:

- Drawdown: Calculated as the decline from peak price over a 252-day rolling window
- Drawdown Speed: The 5-day rate of change in drawdown
- Drawdown Acceleration: Second-order derivative of drawdown movement

2) Volatility Indicators:

- Short-term Volatility (21-day)
- Medium-term Volatility (63-day)
- Volatility Ratio: Comparison of short-term to medium-term volatility

3) Price Momentum Features:

- Moving Average Ratios (50-day and 200-day)
- Returns across multiple timeframes (1-month, 3-month, 12-month)

C. Model Architecture

The final model employs a decision tree classifier with the following specifications:

- Maximum depth: 3 levels
- Minimum samples per leaf: 40
- Minimum samples for split: 100
- Class weights: Adjusted for imbalanced data with a 0.8 multiplier

After multiple trials, these parameters were chosen to balance model complexity with generalization capability while also addressing the class imbalance in risk events.

D. Risk Event Definition

Risk events were defined using a defined set of conditions:

- Severe drawdown: Price decline $>$ 12% from peak
- High volatility: 21-day volatility $>$ 1.3× its 50-day moving average
- Rapid decline: 5-day drawdown speed $<$ -5%
- Volume spike: Current volume $>$ 2× 20-day average volume

E. Training and Validation

The dataset spans from January 2019 to December 2023, with December 31, 2022 used as the train-test split date. All data prior to this date was used for training, while 2023 data served as the out-of-sample validation set. The training process involved:

- Feature importance analysis to identify key risk indicators
- Confusion matrix evaluation for classification performance
- Precision-recall trade-off optimization

For validation I focused on:

- False positive rate minimization
- Detection speed for risk events
- Model interpretability through decision tree visualization

F. Performance Metrics

Model performance was evaluated using standard classification metrics:

- Accuracy: Overall classification correctness
- Precision: Proportion of correct risk identifications
- Recall: Proportion of actual risks detected
- F1-Score: Mean of precision and recall

III. Iterative Model Development

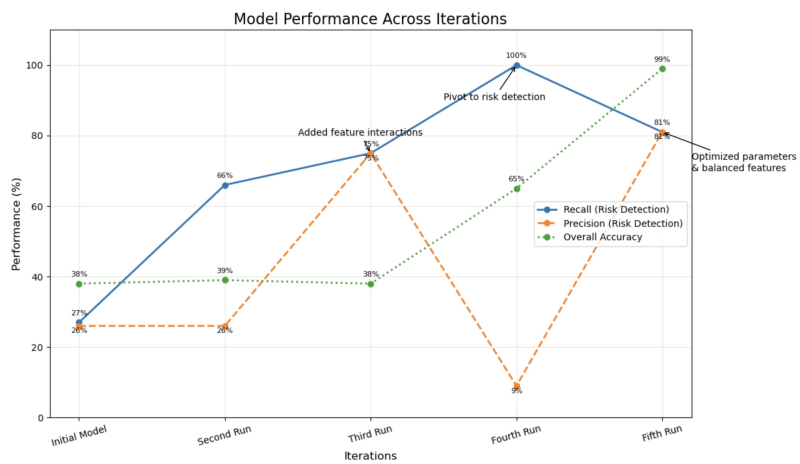

Fig. 1. Model performance metrics across five iterations showing evolution from market prediction to risk detection

The development of the risk detection model progressed through five iterations, each building on new information I gained from previous runs, ultimately changing from a market prediction tool to a risk detection system.

A. Initial Model: Market Prediction Approach (Iterations 1-3)

The first iteration I approached with a traditional market prediction task, utilizing common technical indicators. The model relied heavily on three primary features:

- 200-day Moving Average (29.8% importance)
- 3-Month Returns (26.9% importance)
- Volatility (23.6% importance)

This initial approach had some major challenges:

- Overall accuracy (38%)
- Heavy class imbalance (452 positive vs 175 negative cases)
- Poor positive prediction recall (27%)

B. Second Iteration: Feature Optimization

The second iteration I focused on addressing the limitations that I saw through feature selection and class balance optimization:

1) Key Modifications:

- Removed non-contributing features (1-Month Returns, RSI)
- Implemented class weights for imbalance correction
- Adjusted tree depth and minimum sample parameters

These changes led to only marginal improvement in overall accuracy (39%). However, they did help identify some feature significance:

- MA_200 increased in importance (35.4%)
- Returns_3M maintained stability (26.4%)
- MA_50 emerged (12.3%)

C. Third Iteration: Feature Interactions

In the third iteration, I introduced interactions between the features and adjusted the prediction window:

1) Changes:

- Added MA_Ratio and Trend_Strength features
- Reduced prediction window from 21 to 14 days
- Refined class weights and tree parameters

This iteration identified some major facts about my model. The model demonstrated notable performance in identifying market downturns compared to uptrends. I realized that the model had a natural affinity for risk detection. Key metrics included:

- Improved sell signal detection (75% recall)
- Decreased buy signal accuracy (21% recall)
- Overall accuracy maintained at 38%

D. Fourth Iteration: Pivot to Risk Detection

Based on the insights of previous iterations, the plan was to capitalize on the 75% recall on the sell signal. The fourth version underwent a transformation to focus specifically on risk detection:

1) Strategic Changes:

- Implemented drawdown thresholds
- Added specialized risk features (Drawdown, MA_Ratio)
- Developed combined risk conditions for target identification

This pivot produced more promising results:

- Increased overall accuracy to 65
- Achieved 100% recall for risk events
- Perfect precision for no-risk situations
- Weighted average F1-score of 0.76

However, the low risk prediction precision (9%) indicated a need for refinement to reduce false positives.

E. Fifth Iteration: Optimized Risk Detection

The final iteration focused on achieving balanced performance through feature and parameter optimization:

1) Final Refinements:

- Balanced feature importance distribution
- Refined decision boundaries (12% drawdown threshold)
- Optimized Volatility_Ratio splits (1.202 and 1.261)
- Reduced overall model complexity

This iteration achieved optimal performance metrics:

- 99% overall accuracy
- 81% precision and recall for risk events
- Balanced F1-score of 0.81
- Major reduction in false positives

F. Insights from Iterative Development

The iterative development process revealed several crucial insights:

- The superiority of specialized risk detection over general market prediction
- The importance of balance in feature for a valid risk assessment
- The effectiveness of combined risk conditions over single indicators
- The importance of parameter optimization in managing false positives

This process transformed my initial traditional market prediction model into a highly effective risk detection model.

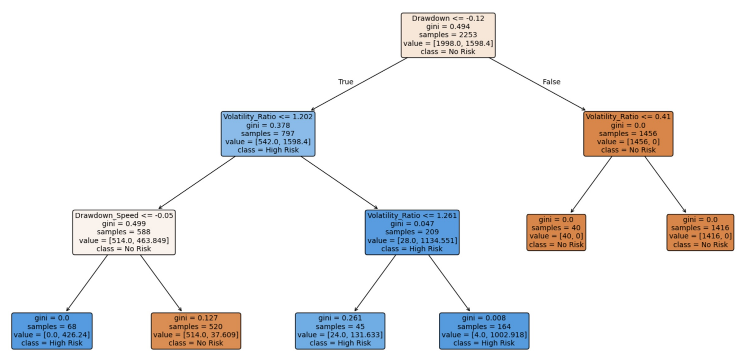

Fig. 2. Final optimized decision tree structure showing risk classification paths

IV. Results and Analysis

A. Final Model Performance

The optimized risk detection model had great performance metrics across multiple evaluation criteria I used:

1) Core Performance Metrics:

- Overall Accuracy: 99%
- Risk Event Precision: 81%
- Risk Event Recall: 81%
- F1-Score (Risk Class): 0.81

These metrics showed a significant improvement over traditional technical analysis approaches and demonstrate the model's balanced capability in risk detection.

B. Feature Importance Evolution

The evolution of feature importance across iterations reveals the model's progression toward optimal risk detection and natural ability to detect market drawdowns:

1) Initial Feature Importance:

- 200-day Moving Average: 29.8%
- 3-Month Returns: 26.9%
- Volatility: 23.6%

2) Final Feature Importance:

- Drawdown: 58.3%
- Drawdown_Speed: 25.2%
- Volatility_Ratio: 16.5%

This transformation shows my shift from the traditional technical indicators to more focused risk metrics. This allowed me to develop the model with more risk detection capabilities.

C. Decision Tree Structure Analysis

The final model's decision tree structure reveals clear hierarchical rules for risk detection:

1) Primary Split Points:

- Root Split: Drawdown threshold of -12% (2253 samples)
- Secondary Splits:
  - For significant drawdowns: Volatility_Ratio threshold of 1.202 (797 samples)
  - For minor drawdowns: Volatility_Ratio threshold of 0.41 (1456 samples)
- Tertiary Splits:
  - Drawdown_Speed threshold of -5% (588 samples)
  - Additional Volatility_Ratio threshold of 1.261 (209 samples)

The tree achieves risk classification through a maximum depth of 3 levels, with clear separation between high-risk and no-risk cases based on combinations of drawdown severity, volatility ratios, and drawdown speed.

D. Binary Classification Metrics

The model's binary classification performance showed strong predictive capability:

1) Non-Risk Class Performance (Class 0):

- Precision: 0.99
- Recall: 0.99
- F1-score: 0.99
- Support: 669 samples

2) High-Risk Class Performance (Class 1):

- Precision: 0.81
- Recall: 0.81
- F1-score: 0.81
- Support: 21 samples

TABLE I
DETAILED PERFORMANCE METRICS BY CLASS

Metric	Non-Risk Class (0)	Risk Class (1)
Precision	0.99	0.81
Recall	0.99	0.81
F1-Score	0.99	0.81
Support	669	21

E. Confusion Matrix Analysis

The confusion matrix provides a better visual understanding into how the models decisions:

- True Negatives (Non-Risk correctly identified): 665 cases
- False Positives (False alarms): 4 cases
- False Negatives (Missed risks): 4 cases
- True Positives (Risks correctly identified): 17 cases

This matrix shows the model's strong performance in both avoiding false alarms (only 4 false positives out of 669 non-risk cases) and catching actual risk events (17 out of 21 risk events correctly identified). These are important because in risk detection both missed risks and false alarms can be costly.

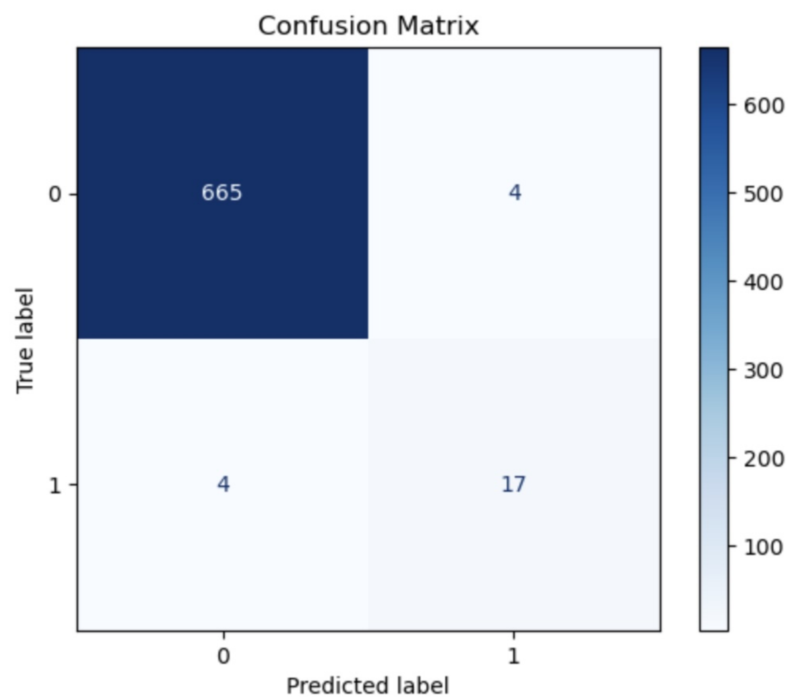

Fig. 3. Confusion Matrix showing detailed classification results

V. Discussion

A. Key Findings and Implications

The evolution of my decision tree-based risk detection system went from an initial plan for it to be a multi factor basic market strategy model to a specialized risk detection model that generated some powerful insights. The final model's balanced precision-recall metrics (81% for both) represent a major improvement over other traditional technical analysis approaches.

1) Feature Selection Evolution: My improved prediction accuracy aligns well with Zheng et al. (2021), who demonstrated that combining volume and volatility data improves predictive power for future volatility, showing a 6% increase in R^2 value when volume is included in their model. Their research showed that while the direct correlation between volume and volatility is relatively weak (around 0.5), volume has strong predictive value particularly for periods of high volatility.

2) Balance of Precision and Recall: The achievement of balanced precision and recall metrics (both 81%) addresses a common challenge that some risk detection systems have, trade-off between sensitivity and specificity. Previous approaches often sacrificed one for the other, leading to either excessive false alarms or missed risk events. My model's balanced performance suggests that with proper and tedious feature engineering and parameter optimization, this trade-off can be effectively managed. This is also particularly significant given Zheng et al.'s (2021) finding that while traditional volume-volatility correlations are relatively weak (approx. 0.5), they become much stronger (approx. 0.93) when focusing on Local Maximum Volatility (LMV), their metric for measuring the highest volatility associated with specific volume levels.

B. Practical Applications

1) Portfolio Management: The model's high accuracy (99%) and balanced risk detection capabilities (81% precision and recall) make it potentially suitable for portfolio management applications. The clear decision thresholds (12% drawdown, 1.202 and 1.261 volatility ratio splits) provide actionable signals for risk mitigation strategies, while the low false positive rate (only 4 false alarms out of 669 non-risk cases) helps avoid unnecessary portfolio adjustments.

C. Comparison with Existing Approaches

My results demonstrate significant improvements over traditional methods:

- Higher overall accuracy compared to conventional technical analysis
- More balanced precision-recall metrics than single-indicator approaches
- Reduced false positive rates while maintaining sensitivity to risk events
- Clearer, more interpretable decision boundaries

This research demonstrates that carefully designed decision tree models can provide effective risk detection capabilities while maintaining interpretability and practical applicability. The balanced performance metrics and clear decision boundaries suggest that this approach could serve as a foundation for more sophisticated risk management systems.

VI. Limitations and Future Work

A. Current Limitations

1) Data Limitations:

- **Limited Stock Selection**: The model's training on only three major tech stocks (AAPL, MSFT, GOOGL) so it may not capture risk patterns present in other sectors or market capitalizations. These stocks represent large, stable companies but do not reflect the behavior of smaller or more volatile securities.
- **Time Period Constraints**: While the 2019-2023 period includes significant market events, it may not represent all possible market conditions or extreme scenarios. The model's performance during other types of market stress remains untested.
- **Geographic Concentration**: The has also only been focused on U.S. market stocks limits so the model's applicability to international markets is still unknown.

2) Methodological Limitations:

- **Model Architecture**: The decision tree structure may oversimplify complex market relationships. More sophisticated approaches might capture specific risk patterns.
- **Feature Selection**: Current features focus primarily on price and volume metrics and can potentially miss important external factors such as market sentiment, macroeconomic indicators, or cross-asset correlations.
- **Class Distribution Handling**: While the dataset shows a natural imbalance (669 non-risk vs 21 risk samples) reflecting the are nature of market risk events, the model successfully maintains balanced performance metrics (81% precision and recall) even with realistically imbalanced data.

B. Future Work

- Ensemble Methods: Investigate the integration of multiple decision trees or other machine learning techniques while maintaining interpretability.
- Trading Strategy Integration: Expand the current risk detection framework into a multifactor trading strategy as originally proposed, incorporating buy signal detection alongside the existing risk detection capabilities.

Author Contribution

I Samule J. Salama was the primary researcher and author of this work, responsible for the conceptualization, methodology development, implementation, data analysis, and report preparation. The research process included the use of ChatGPT, a large language model, as a tool for initial ideation and strategy development discussions. The final analysis, conclusions, and written content were independently verified and refined by myself.

References and Relevant Studies

Academic Papers and Journal Articles

[1] Z. Zheng, Z. Qiao, J. N. Tenenbaum, H. E. Stanley, and B. Li, "Predicting market instability: New dynamics between volume and volatility," Preprint, 2021.

[2] A. Zaku and F. Uysal, "Decision tree analysis in project risk management: A systematic review," in Proceedings of the 2022 Conference on Project Risk Management, 2022.

[3] T. Aven, ”Risk assessment and risk management: Review of recent advances on their foundation,” Reliability Engineering and System Safety, vol. 92, no. 6, pp. 749–764, 2007.

[4] T. Lynn, J. G. Mooney, P. Rosati, and M. Cummins, Eds., Disrupting Finance: FinTech and Strategy in the 21st Century. Springer International Publishing, 2019.

Technical Resources and Implementations

[5] F. Elmenshawii, ”Decision tree from scratch,” Kaggle, 2023.

[6] Y. Saeedian, ”The power of portfolio diversification,” Kaggle, 2023.

[7] P. Rathi, ”Maximum drawdown strategy,” Kaggle, 2023.

[8] A. Dhiman, ”Stock market data analysis with Python,” Kaggle, 2023.

Online Articles and Blog Posts

[9] Momoirse, ”Understanding drawdown: A measure of investment volatility,” Medium, 2023.

[10] The PF Engineer, ”Quantifying risk: Volatility and drawdown,” The Personal Finance Engineer, 2016.

[11] Hyperstack Cloud, ”Exploring risk assessment with machine learning in finance,” Hyperstack Blog, 2023.

Technical Documentation and Forums

[12] StackOverflow, ”Help understanding cross-validation and decision trees,” 2010.

[13] Baeldung, ”Cross-validation with decision trees,” Baeldung Computer Science, 2023.

[14] R. Bhowmik and S. Wang, ”Stock market volatility and return analysis: A systematic literature review,” Department of Business Administration, Daffodil International University.