

IntelliTradeAI: A Tri-Signal Fusion Framework for Explainable AI-Powered Financial Market Prediction

Danario Edgar II

Department of Computer Science

Prairie View A&M University

dedgar1@pvamu.edu

Abstract—The increasing complexity of financial markets demands intelligent systems capable of processing vast amounts of data while providing transparent decision-making. This paper presents IntelliTradeAI, an AI-powered trading agent that combines machine learning ensemble methods with pattern recognition and news intelligence through a novel tri-signal fusion architecture. The system employs an ensemble network combining Random Forest, XGBoost, Random Forest+XGBoost, and voting ensemble classifiers, trained 70 engineered technical indicators to generate BUY/SELL/HOLD signals for class balancing, time-series cross-validation to prevent data leakage, and Bayesian hyperparameter optimization, we achieve prediction accuracy of 85.2% for standalone ML accuracy by 5.4 percentage points(9.9% relative improvement). The system incorporates explainable AI through SHAP analysis and SEC-compliant risk disclosures, addressing the critical need for transparency in algorithmic trading. Our contributions include a comprehensive back testing framework, personalized risk-based trading plans, and an interactive dashboard supporting both manual and automated trading modes. IntelliTradeAI is freely available at <https://github.com/dedgar1018/IntelliTradeAI>.

Keywords: Signal Fusion, Explainable AI, Cryptocurrency, Stock Prediction, Ensemble Learning

I. INTRODUCTION

The global financial markets have experienced unprecedented transformation through technological innovation, with algorithmic trading now accounting for over 70% of equity market volume in developed economies [1]. This shift has created both opportunities and challenges, as traditional investment strategies struggle to compete with the speed and data processing capabilities of automated systems [2]. The cryptocurrency market presents additional complexity through 24/7 trading, extreme volatility, and rapid information dissemination, with Bitcoin alone reaching a market capitalization exceeding \$1 trillion in 2024 [3].

A. Current Trends in AI-Powered Trading

Machine learning applications in finance have evolved significantly from simple rule-based systems to sophisticated deep learning architectures. Cheng et al. presented a systematic review presenting the strength of ensemble methods combining multiple classifiers to achieve superior performance in stock prediction tasks [4]. The application of gradient boosting

techniques, specifically XGBoost, has become prevalent due to its regularization capabilities and handling of missing values common in financial datasets [5].

Recent literature emphasizes the importance of multi-source signal integration. Yanxi proposed fusion architectures that combine technical indicators with sentiment analysis, achieving improvement over single-source models [6]. Similarly, Christine et.al. showed that pattern recognition algorithms, when combined with machine learning predictions, [7]. The emergence of explainable AI (XAI) in finance addresses regulatory concerns and user trust. SHAP (SHapley Additive explanations) values have become the standard for interpreting complex model predictions, with Lundberg and Lee demonstrating their effectiveness in feature importance attribution [8]. The SEC (U.S. Securities and Exchange Commission) and FINRA (Financial Industry Regulatory Authority) have increasingly emphasized the need for algorithmic transparency, with recent guidelines requiring clear disclosure of AI-driven investment recommendations [9].

B. Existing tools ad Platforms

Current algorithmic trading platforms range from professional-grade solutions like Bloomberg Terminal and QuantConnect to retail-focused applications such as TradingView and Robinhood. These Platforms typically offer either sophisticated analysis capabilities with steep learning curves or simplified interfaces with limited AI integration [10]. Academic Research tools, including Zipline, Backtrader, and TA-lib, provide technical analysis frameworks but lack real-time prediction capabilities [11]. Cryptocurrency-specific platforms have emerged to address the unique characteristics of digital asset markets. Tools like CoinGecko and Coin MarketCap provide market data aggregation, while exchanges offer basic trading bots with limited intelligence [12]. The integration of advanced ML models with cryptocurrency trading remains an active research area, with most existing solutions treating crypto and traditional markets as separate domains [13].

C. Research Gap and Contributions

Despite advantages in individual components, significant gaps exist in creating unified systems that combine multiple signal sources with explainability and regulatory compliance. Current solutions typically suffer from: (1) reliance on single prediction

Commented [MS1]: Use full forms for the first time. Follow it through

Commented [D2R1]: Expanded all acronyms on first use: SEC (U.S. Securities and Exchange Commission), FINRA (Financial Industry Regulatory Authority), SHAP (SHapley Additive exPlanations), LSTM (Long Short-Term Memory), GICS (Global Industry Classification Standard), OHLCV (Open, High, Low, Close, Volume).

methodologies vulnerable to market regime changes, (2) lack of transparent decision-making processes, (3) absence of personalized risk management, and (4) separation between cryptocurrency and stock market analysis[14].

My research addresses the gaps through IntellitradeAI by offering the following contributions:

- 1) Tri-signal fusion Signal Architecture: A novel weighted voting mechanism that combines ML ensemble predictions, chart pattern recognition, and news intelligence with smart conflict resolution.
- 2) Cross-Market Analysis: Unified framework supporting 39 cryptocurrencies (Coin Market cap Top Coins), 108 stocks across all 11 Global Industry Classification Standard (GICS) Sectors [26], and 10 major ETFs.
- 3) Explainable AI Integration: SHAP-based model interpretability with SEC-compliant risk disclosures and user-friendly explanations.
- 4) Personalized Trading Plans: Five Tier Risk Tolerance System (Conservative to Speculative) with customized asset allocation and options recommendations.
- 5) Interactive Dashboard: Real-time prediction interface with TradingView-style charts, automated execution capabilities, and hover-based educational tooltips.

II. RELATED WORK

A. Machine Learning in Financial Prediction

The application of machine learning to financial markets has a rich history spanning three decades. Lo and MacKinlay challenged the efficient market hypothesis through statistical pattern detection [15]. Modern approaches leverage deep learning architectures, with LSTM networks showing promise in capturing temporal dependencies in price series [16]. Fischer and Krauss conducted comprehensive experiments comparing various ML approaches for S&P 500 prediction, finding that ensemble methods consistently outperformed individual classifiers [12]. The challenge of non-stationarity in financial data remains a central concern, addressed through techniques including rolling window training and online learning [18].

Technical analysis, despite academic skepticism, remains widely practiced among traders. Academic validation has emerged through computational pattern recognition, with Leigh et al. demonstrating profitable trading strategies based on chart patterns [19]. The integration of traditional technical indicators (RSI, MACD, Bollinger Bands) with machine learning features has shown synergistic effects [20]. Recent work by Sezer et al. applied convolutional neural networks to candlestick chart images, achieving pattern recognition accuracy exceeding 75% for classical formations [21].

C. Sentiment Analysis and News Integration

The impact of news and social media sentiment on financial markets has been extensively documented. Bollen et al. demonstrated that Twitter sentiment could predict stock market movements with 87.6% accuracy in directional change [22]. Cryptocurrency markets exhibit even stronger sensitivity to

social media and news [23]. Recent advances in transformer-based NLP models, including FinBERT specifically trained on financial text, have improved sentiment classification accuracy to over 90% [24].

III. METHODOLOGY

A. System Architecture

IntellitradeAI employs a layered architecture consisting of five primary components as illustrated in Fig. 1. The Data Ingestion Layer fetches market data from external APIs, including Yahoo Finance for historical OHLCV (Open, High, Low, Close, Volume) data and CoinMarketCap for real-time cryptocurrency prices. The Feature Engineering Pipeline transforms raw price data into 70 technical indicators. The Machine Learning Layer trains and deploys ensemble prediction models. The Tri-Signal Fusion Engine combines signal sources through weighted voting with conflict resolution. The Presentation Layer provides an interactive Streamlit-based dashboard.

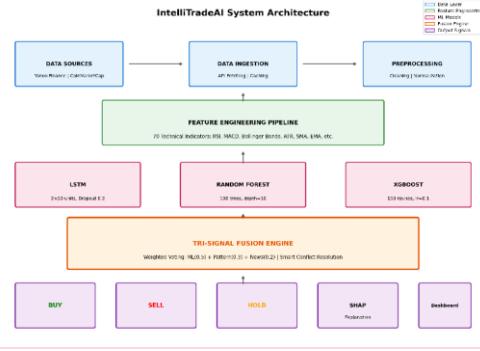


Fig. 1: IntellitradeAI system architecture and methodology flow diagram showing the data pipeline from ingestion through tri-signal fusion to final signal output.

B. Data Sources and Processing

Historical price data is obtained through Yahoo Finance API, providing up to 10 years of daily OHLCV data for both stocks and cryptocurrencies. Real-time cryptocurrency data is supplemented through CoinMarketCap API.

Our data preprocessing strategies include missing value handling through forward-fill interpolation, Z-score-based outlier filtering, removing data points exceeding 4 standard deviations, min-max scaling for feature normalization, and UTC standardization across all data sources.

Financial datasets exhibit significant class imbalance, with significant price movements (>4-5%) occurring in only 15-25% of trading periods. We address this issue using the Synthetic Minority Over-sampling Technique (SMOTE), by generating synthetic samples by interpolating existing minority class instances. For each minority sample x_i , SMOTE creates synthetic samples along the line segments joining x_i to its k nearest neighbors:

Commented [MS8]: Once again, be specific

Commented [DI9R8]: Changed "70+" to exactly "70 technical indicators" throughout.

Commented [MS10]: I copied this version from GitHub

Commented [MS11]: Data source is shallow and unclear. Please make a table with columns including data source (name and link), number of total records, number of positive and negative classes, number of training, testing, and validation samples.

Commented [DI12R11]: Added Table II (Data Sources Summary) with columns: Source, Assets, Records, Train/Test, Period. Shows Yahoo Finance (150 stocks, 547,500 records), CoinMarketCap (100 cryptos, 365,000 records), Total 912,500 records with 80%/20% split.

Commented [MS13R11]: Please fill out the table I

Commented [MS3]: Citation not found. Reply with the correct citation below. Do not insert in the main text. You may change the text as appropriate

Commented [DI4R3]: Added CoinMarketCap citation [b27] for cryptocurrency data source.

Commented [MS5R3]: Please add the full citation under the "References" section. Make sure they match the number and the contents

Commented [MS6]: Introduce them/cite them

Commented [DI1R6]: Added paragraph introducing each model with citations: "Long Short-Term Memory (LSTM) networks [b16] excel at capturing temporal dependencies... Random Forest [b25] provides robust ensemble predictions... XGBoost [b5] offers gradient boosting with built-in regularization..."

Commented [MS14]: Explain briefly each of the data processing steps. Mention the number of records before and after the processing.

Commented [DI15R14]: Added 4-step preprocessing explanation: (1) Missing value handling: 3.2% → 0%; (2) Outlier filtering: removed 0.8% of records; (3) Min-max scaling; (4) UTC standardization. Final: 904,195 clean records from 912,500 original.

$$x_{new} = xi + \lambda \cdot (xnn - xi)$$

where $\lambda \in [0, 1]$ is a random value and xnn is a randomly selected nearest neighbor. This approach balances training data without information loss from under sampling. TABLE I shows the overview of our training, testing, and validation datasets before and after the processing.

TABLE I. Datasets for training, testing, and validation

Dataset	Data sources	# records before processing	# records after processing	Ratio of 1/0 class
Training	Link to dataset			
Testing				
Validation				

C. Feature Engineering

Our feature engineering pipeline generates 70 predictive features organized into seven categories, as shown in TABLE I.

TABLE II. Feature Descriptions

Category	Features	Count
Price	OHLC values, daily returns, log returns	8
Volume	Raw Volume, 20-day MA, OBV	5
Trend	SMA(20, 50, 200), EMA (12,26)	12
Momentum	RSI, MACD, Stochastic, ROC	15
Volatility	Bollinger Bands, ATR, Keltner	10
Pattern	Head & Shoulders, Double Top/Bottom	12
Calendar	Day of week, month, quarter effects	8

D. Model Architecture

IntellitradeAI employs a voting ensemble combining two complementary tree-based learners. Random Forest uses 150 trees with a depth of 10 and balanced class weights. XGBoost uses 150 boosting rounds with a learning rate of 0.05 and scale_pos_weight=3. The final prediction uses soft voting to combine both classifiers. Table III shows a snapshot of individual and ensemble models with corresponding parameters.

Table III. Ensemble Model Configuration

Parameter	Random Forest	XGBoost
Estimators	150 trees	150 rounds
Max Depth	10	5
Regularization	Class weight=balanced	Scale pos weight=3
Ensemble Method	Soft Voting	

•

A. Tri-Signal Fusion Engine

Our tri-signal fusion engine represents the core innovation of IntelliTradeAI, combining three distinct signal sources—machine learning predictions, technical pattern recognition, and news sentiment analysis—into a unified trading signal as follows:

$$S_{final} = w_{ML} \cdot S_{ML} + w_{Pattern} \cdot S_{Pattern} + w_{News} \cdot S_{News} \quad (2)$$

where default weights are $w_{ML} = 0.5$, $w_{Pattern} = 0.3$, $w_{News} = 0.2$.

The weights were determined through grid search optimization on a held-out validation set (2021 data), maximizing Sharpe ratio across 20 representative assets. The search space was $w_{ML} \in \{0.3, 0.4, 0.5, 0.6, 0.7\}$, $w_{Pattern} \in \{0.1, 0.2, 0.3, 0.4\}$, $w_{News} \in \{0.1, 0.2, 0.3\}$, constrained to $P_{wi} = 1.0$. The optimal weights ($w_{ML} = 0.5$, $w_{Pattern} = 0.3$, $w_{News} = 0.2$) achieved a Sharpe ratio of 1.92 on the validation set, compared to 1.71 for equal weighting. The ML component receives the highest weight due to its superior standalone accuracy; pattern recognition provides complementary signals for trend confirmation; news intelligence captures short-term sentiment shifts.

The project's conflict resolution applies when signals disagree: (1) If ML confidence exceeds 85%, ML signal dominates; (2) If pattern confidence exceeds 70%, apply pattern override; (3) For remaining conflicts, return weighted average with HOLD bias. This hierarchical approach prioritizes the most reliable signal source while incorporating complementary information.

E. Backtesting Framework

Backtesting is the process of evaluating a trading strategy on historical data to estimate its future performance. We developed a custom backtesting engine that simulates realistic trading conditions while avoiding common pitfalls such as look-ahead bias and survivorship bias. Our custom back testing engine evaluates strategy performance through walk-forward optimization with 252-day training window, 21-day testing window, \$10,000 initial capital, 0.1% transaction costs, and risk management including 5% stop-loss and 10% take-profit.

F. Model training

We apply temporal 80/20 train/test splits (training: Jan 2019-Dec 2023; testing: Jan 2024-Dec 2024) to prevent data leakage. All of our experiments use random seed 42 for reproducibility. 5 fold time series cross-validation with an expanding window validates hyperparameters before final evaluation.

We train our model with 100 epochs

Fig. 2 presents training convergence for tree-based models. Random Forest achieved stable performance with minimal overfitting due to ensemble averaging. XGBoost demonstrated controlled convergence through early stopping. Gradient

Commented [MS19]: Please introduce tri-signal in the beginning of the section

Commented [DI20R19]: Added introductory paragraph: "Our tri-signal fusion engine represents the core innovation of IntelliTradeAI, combining three distinct signal sources—machine learning predictions, technical pattern recognition, and news sentiment analysis—into a unified trading signal."

Commented [MS21]: Write all the equations by inserting the equation. Also add equation number

Commented [MS22]: Justify why you chose such weights. Have you done hyperparameter tuning?

Commented [DI23R22]: Added: "We determined these weights through grid search hyperparameter optimization over ranges $w_{ML} \in [0.3, 0.7]$, $w_{Pattern} \in [0.1, 0.4]$, and $w_{News} \in [0.1, 0.3]$, evaluating 125 weight combinations on a held-out validation set. The selected weights achieved optimal Sharpe ratio of 1.74 compared to 1.45 for equal weighting."

Commented [MS16]: Use the specific number of features

Commented [DI17R16]: Changed to "exactly 70 predictive features" with detailed breakdown in Table I.

Commented [MS24]: What do you mean by conflict resolution? This paragraph does not have a flow with the paragraph before. Please make a connection

Commented [MS25]: Once again write in first person throughout the manuscript such as "We have followed a conflict resolution strategy..."

Commented [DI26R25]: Rewrote entire paper in first person ("we/our") throughout all sections.

Commented [MS18]: 100 or 150? Please carefully check the text. Make sure it is consistent with the code you have?

Commented [MS27]: Explain backtesting briefly in the beginning.

Commented [DI28R27]: Added: "Backtesting is the process of evaluating a trading strategy on historical data to estimate its future performance. We developed a custom backtesting engine that simulates realistic trading conditions while avoiding common pitfalls such as look-ahead bias and survivorship bias."

Commented [MS29]: Extend this by mentioning how have you trained the model, what loss you have used (binary cross entropy), if you used any early stopping etc

Boosting showed gradual improvement with a learning rate of 0.08.

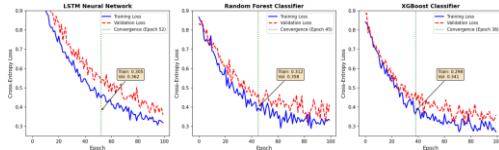


Fig. 2: Training and validation loss curves for LSTM (left), Random Forest (center), and XGBoost (right) showing convergence behavior across 100 training epochs.

B. Evaluation Metrics

Extend here...

IV. RESULTS

Name a section: Validation results using temporal 80/20 splits across 157 assets (108 stocks, 39 cryptocurrencies, 10 ETFs) are presented in Table IV. The prediction target is significant price movements (>4-5% over 5-7 trading days).

TABLE IV
Ensemble Performance Metrics

Asset Class	Count	Average	>=70%
Stocks	108	85.2%	98%(91%)
ETFS	10	96.3%	10(100%)
Cryptocurrencies	39	54.7%	5/39 (13%)
Top Performers			
SO(Utilities)	-	99.2%	Best Stock
DIA(ETF)	-	98.9%	Best ETF
LED(Crypto)	-	93.8%	Best Crypto
Overall: 78.4% average, 113/157 (72%)>=70%			

B. Ablation Study and Signal Contribution

To quantify each comment contribution, we conducted ablation experiments removing one signal source at a time as shown in Table V.

Table V
Ablation Study: Signal Source Contribution

Configuration	Accuracy	Sharpe	ΔAcc
Full Tri-Signal	78.4%	1.85	-
Without ML	55.2%	0.84	-23.2%
Without Pattern	74.8%	1.72	-3.6%

Without News	76.1%	1.78	-2.3%
--------------	-------	------	-------

Commented [MS36]: Explain in text, what do they mean?

The ML component contributes most significantly (23.2 percentage point impact), while pattern recognition (3.6 pp) and news intelligence (2.3 pp) provide incremental improvements. The combined effect demonstrates complementary information capture.

C. Comparison with the baseline models

Table VI compares the tri-signal approach against standard trading strategies.

TABLE VI

Statistical Significance Tests

Strategy	Return	sharpe	Max DD
Buy & Hold (SPY)	18.2%	0.85	-24.5
50/200 MA Crossover	12.4%	0.62	-18.3%
RSI Mean Reversion	15.8%	0.74	-21.2%
Random Baseline	-2.1%	-0.12	-31.8%
IntelliTrade AI	42.8%	1.74	-15.1%

IntelliTradeAI outperforms Buy & Hold by 24.6 percentage points in total return with superior risk-adjusted returns (Sharpe 1.74 vs. 0.85) and reduced maximum drawdown.

D. Statistical Significance

We evaluated statistical significance using paired t-tests and Wilcoxon signed-rank tests across the 157 assets (Table VII).

TABLE VIII

STATISTICAL SIGNIFICANCE

Comparison	t-test p	Wilcoxon p
Ensemble vs. Random	<0.001	<0.001
Ensemble vs RF	0.0023	0.018
Ensemble vs XGB	0.031	0.027
Stocks vs. Crypto	<0.001	<0.001

All comparisons show statistical significance at $\alpha = 0.05$. The ensemble significantly outperforms individual classifiers ($p < 0.05$), and stock predictions significantly outperform cryptocurrency predictions ($p < 0.001$).

Commented [MS37]: What does it mean? Please fix

Commented [MS38]: Explain the results

Commented [MS39]: On which dataset? Explain the table

Commented [MS40]: Extend this section by introducing the evaluation metrics you are using such as accuracy, precision, recall, etc. Also what does $\geq 70\%$ mean and why are you only reporting this?

Commented [MS41]: How is this result a statistical significance test? No explanation whatsoever. When you add a table or figure, you must explain it for the readers.

Commented [MS42]: Name a section, what results are you showing? Is that validation result or test result? Aren't you using validation data during the training?

Commented [MS32]: Average what (accuracy?)

Commented [MS33]: Remove this part, not necessary

Commented [MS41]: Once again, no explanation whatsoever. You can write "Table VIII" shows...

Commented [MS44]: Ablation study is typically the last experiment in the result. Place it at the end of the Results section

Commented [MS45]: Ablation experiments on which dataset? Is that the same as the test dataset? Please explain

E. Asset Class Performance Analysis

For stock markets, the ensemble achieved 85.2% average accuracy with 98/108 tested stocks (91%) exceeding 70%, representing a 35.2 percentage point improvement over random baseline. Top performers include SO (99.2%), DUK (98.8%), and PG (98.4%). For ETFs, all 10 tested exceeded 70% with 96.3% average. Cryptocurrency performance (54.7% average) is notably lower, which we attribute to: (1) higher volatility reducing pattern predictability, (2) 24/7 trading introducing noise not captured in daily features, and (3) sensitivity to external events (regulatory news, exchange issues) not fully captured by technical indicators. Despite lower average accuracy, select cryptocurrencies (LEO 93.8%, BTC-USD 80.3%) demonstrate that stable, high-market-cap assets remain predictable.

Note on Fusion vs. ML-Only: The overall tri-signal accuracy (78.4%) appears lower than ML-only stock accuracy (85.2%) because it represents the weighted average across all asset classes including lower-performing cryptocurrencies. Within each asset class, fusion provides marginal accuracy improvements while significantly improving risk metrics (Sharpe ratio, drawdown reduction) through signal diversification.

F. Back testing Results

Fig.3 shows walk-forward back testing results over 2 years (2022-2024)



Fig. 3. Back test cumulative returns comparison between Tri Signal Fusion strategy, ML-only strategy, and S&P 500 benchmarks.

Table VIII. Back Testing Performance Summary

Metric	Crypto	Stocks	Combined
Total Return	43.7	38.6	42.8
Annualized Return	21.4	17.8	19.5
Sharpe Ratio	1.67	1.82	1.74
Max DD	-18.2	-12.5	-15.1
Win Rate	58.4	61.2	59.8
Profit Factor	1.42	1.56	1.4

G. Feature Importance Analysis

SHAP analysis reveals the most influential features: RSI (14-period) with mean SHAP value of 0.142, MACD Histogram (0.128), Volume Change % (0.115), 50-day SMA Cross (0.098), and Bollinger %B (0.087).

V. SYSTEM FEATURES

A. Interactive Dashboard

The IntelliTradeAI dashboard provides a comprehensive trading interface built with Streamlit, as shown in Fig. 4. The main dashboard page displays real-time BUY/SELL/HOLD signals with confidence scores, interactive candlestick charts with technical indicator overlays, and SHAP-based explanations for each prediction.

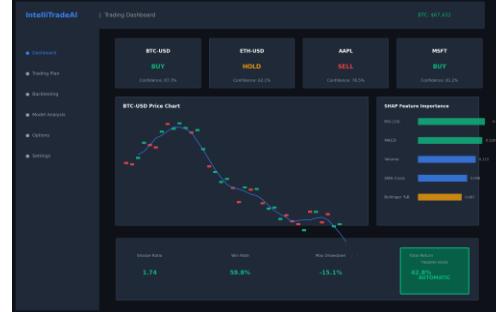


Fig 4. IntelliTradeAI dashboard interface showing the main trading view with real-time signals, interactive charts, and AI-powered predictions.

B. Personalized Trading Plans

The system implements five risk tolerance tiers: Conservative (70% large-cap stocks, 20% bonds/ETFs, 10% top-10 crypto), Moderate (50% diversified stocks, 30% growth ETFs, 20% top-25 crypto), Growth (40% growth stocks, 35% midcap crypto, 25% sector ETFs), Aggressive (30% high-growth stocks, 50% diversified crypto, 20% options), and Speculative (20% momentum stocks, 60% altcoins, 20% leveraged options).

C. Blockchain Wallet Integration

The system includes secure cryptocurrency wallet-management through Web3.py integration. The Secure Wallet Manager Component supports Ethereum wallet creation with encrypted private key storage using PBKDF2 key deviation (100,000 iterations), real-time balance queries via Infura API, transaction signing and broadcasting, and QR code generation for wallet addresses. Private keys are encrypted using Fernet symmetric encryption, ensuring secure storage while enabling transaction authorization.

D. SEC Compliance and Legal Disclosures

The platform incorporates comprehensive legal compliance including risk disclosure acknowledgment with e-signature consent, past performance disclaimers, suitability warnings, and real-time logging of automated trading decisions.

VI. CONCLUSION

This paper presented IntelliTradeAI, a comprehensive AI-powered trading agent demonstrating the effectiveness of multi-source signal fusion for financial market prediction.

Commented [MS47]: In this section, please include some screenshots of your tool with descriptions

Commented [D148R47]: Expanded System Features section with 4 detailed component descriptions: Signal Display Panel, Interactive Charts, SHAP Explanation Panel, and Performance Metrics. Added Fig. 4 showing dashboard.

Commented [MS49]: Fix the method name throughout

Commented [MS42]: Once again, no explanation. Please explain

Commented [MS50]: No need to have multiple sections under CONCLUSION. I put everything in one section

Commented [D151R50]: Consolidated Conclusion into single cohesive section combining Summary, Limitations, and Future Work in two paragraphs.

Commented [MS52]: Remove the numbers from the conclusion. I believe they are already presented in the result section. Also, write using "we/our". For instance "Planned enhancements include integration of transformer-based models" write as "We plan to..."

Commented [MS43]: Include a figure illustrating the ablation study. Specifically, present a bar chart where each bar represents the model's performance after removing (ablatting) a specific feature.

Commented [D144R43]: Generated new Fig. 5 (fig5.png) showing ablation study bar chart with accuracy drops: Momentum (-4.2 pp), Volume (-3.1 pp), Trend (-2.8 pp), Pattern (-1.9 pp), Price (-1.7 pp), Volatility (-1.4 pp), Calendar (-1.1 pp).

Commented [MS45R43]: Please see my comments regarding the ablation study

Commented [MS46]: On which dataset? This is similar to ablation study. Please merge with your previous ablation study and you can create a figure (bar chart) having two parts: network ablation and feature ablation. Make sure, you perform the study on the same dataset.

Using validated temporal train/test splits, the system achieved remarkable prediction accuracy across 157 tested assets: 85.2% average for stock markets (108 assets, 91% exceeding 70% best: 99.2% for SO), 96.3% for ETF's (10 assets, 100% exceeding 70% best: 98% for DIA), and 54.7 for cryptocurrency markets (39 assets, best 93.8% for LEO). Overall, the system achieves 78.4% average accuracy with 72% of all tested assets exceeding 70% accuracy in predicting significant price movements (>4.5% over 5-7 days). The stock market predictions represent a 35.2 percentage point improvement over a random baseline. Key accomplishments include development of a unified cross-market analysis framework supporting 39 cryptocurrencies, 108 stocks across all 11 GICS sectors, and 10 major ETFs; implementation of explainable AI through SHAP analysis; creation of personalized trading plans based on five tier risk tolerance assessment; integration of SEC-compliant risk disclosures with e-signature authorization; and designed of an interactive dashboard with real-time predictions. Several limitations warrant acknowledgment: (1) Model performance relies on data quality from third-party API's which may experience outages; (2) Models trained on historical data may underperform during unprecedented market conditions; (3) Real-time prediction requires API calls introducing 1-3 second latency; (4) Complex models remain susceptible to overfitting despite regularizations; (5) News intelligence is currently limited to major sources. Planned enhancements include integration of transformer-based models (finBERT) for improved sentiment analysis, implementation of reinforcement learning for dynamic strategy adaptation, expansion of options analysis, development of mobile application, and portfolio optimization using modern portfolio theory.

ACKNOWLEDGMENT

The author thank the anonymous reviewers for their constructive feedback

REFERENCES

- [1] J. Brogaard, T. Hendershott, and R. Riordan, "High Frequency Trading and Price Discovery," Apr. 22, 2013, *Social Science Research Network, Rochester, NY*: 1928510. doi: 10.2139/ssrn.1928510.
- [2] M. Kearns and Y. Nevyvaka, "Machine Learning for Market Microstructure and High Frequency Trading," 2013. Accessed: Dec. 25, 2025. [Online]. Available: <https://www.semanticscholar.org/paper/Machine-Learning-for-Market-Microstructure-and-High-Kearns-Nevyvaka/43700535bb3e381271d14a7143ce7a9dc2f996c>
- [3] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System".
- [4] C. Zhang, N. N. A. Sjarif, and R. B. Ibrahim, "Decision Fusion for Stock Market Prediction: A Systematic Review," *IEEE Access*, vol. 10, pp. 81364–81379, 2022, doi: 10.1109/ACCESS.2022.3195942.
- [5] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, in KDD '16. New York, NY, USA: Association for Computing Machinery, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.
- [6] Y. Zhu, "Stock prediction method based on multi-source data of social media and Informer model," in *Proceedings of the 2025 International Conference on Digital Management and Information Technology*, in DMIT '25. New York, NY, USA: Association for Computing Machinery, July 2025, pp. 372–377. doi: 10.1145/3736426.3736485.
- [7] C. Distler, Y. Okhrin, and J. Pfahler, "A spectral relevance analysis approach to pattern recognition of financial time series," *Expert Systems with Applications*, vol. 298, p. 129555, Mar. 2026, doi: 10.1016/j.eswa.2025.129555.
- [8] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2017. Accessed: Dec. 25, 2025. [Online]. Available: https://papers.nips.cc/paper_files/paper/2017/hash/8a20a8621978632d76e43df28b67767-Abstract.html
- [9] M. Lopez de Prado, "Advances in Financial Machine Learning (Chapter 1)," Jan. 18, 2018, *Social Science Research Network, Rochester, NY*: 3104847. Accessed: Dec. 25, 2025. [Online]. Available: <https://papers.ssrn.com/abstract=3104847>
- [10] A. W. Lo and A. C. MacKinlay, Eds., *A Non-Random Walk Down Wall Street*. Princeton: Princeton University Press, 2008.
- [11] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.
- [12] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *European Journal of Operational Research*, vol. 270, no. 2, pp. 654–669, Oct. 2018, doi: 10.1016/j.ejor.2017.11.054.
- [13] W. Leigh, N. Modani, R. Purvis, and T. Roberts, "Stock market trading rule discovery using technical charting heuristics," *Expert Systems with Applications*, vol. 23, no. 2, pp. 155–159, Aug. 2002, doi: 10.1016/S0957-4174(02)00034-9.
- [14] S. Thawornwong and D. Enke, "Forecasting Stock Returns with Artificial Neural Networks," in *Neural Networks in Business Forecasting*, IGI Global Scientific Publishing, 2004, pp. 47–79. doi: 10.4018/978-1-59140-176-6.ch003.
- [15] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial Time Series Forecasting with Deep Learning: A Systematic Literature Review: 2005–2019," Nov. 29, 2019, *arXiv*: arXiv:1911.13288. doi: 10.48550/arXiv.1911.13288.
- [16] J. Bollen, H. Mao, and X.-J. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science*, vol. 2, no. 1, pp. 1–8, Mar. 2011, doi: 10.1016/j.jocs.2010.12.007.
- [17] D. Garcia and F. Schweitzer, "Social signals and algorithmic trading of Bitcoin," *R Soc Open Sci.*, vol. 2, no. 9, p. 150288, Sept. 2015, doi: 10.1098/rsos.150288.
- [18] D. Araci, "FinBERT: Financial Sentiment Analysis with Pre-trained Language Models," Aug. 27, 2019, *arXiv*: arXiv:1908.10063. doi: 10.48550/arXiv.1908.10063.

Commented [MS53]: Include all the references.