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 growing complexity of financial markets requires intelligent systems that can process vast amounts of data while providing transparent decision-making. Here we presents IntelliTradeAI, an Artificial Intelligence (AI)-powered trading agent that combines Machine Learning (ML) ensemble methods with pattern recognition and news intelligence through a novel tri-signal fusion architecture. Our system employs an ensemble framework by combining Random Forest, XGBoost, and voting ensemble classifiers, trained on 70 engineered technical indicators to generate BUY/SELL/HOLD signals. Through Synthetic Minority Over-sampling Technique (SMOTE) class balancing, time-series cross-validation to prevent data leakage, and Bayesian hyperparameter optimization, we achieve prediction accuracy of 85.2% for stocks and 72.9% for the top 10 cryptocurrencies using volatility-aware adaptive thresholds. Our system incorporates explainable AI through SHapley Additive exPlanations (SHAP) analysis and U.S. Securities and Exchange Commission (SEC)-compliant risk disclosures, addressing the critical need for transparency in algorithmic trading. Our contributions include a comprehensive backtesting 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/djedgar1018/IntelliTradeAI.

Index Terms—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 (ML) applications in finance have evolved significantly from simple rule-based systems to sophisticated deep learning architectures. Cheng et al. presented a systematic review demonstrating the strength of ensemble methods combining multiple classifiers to achieve superior performance in stock prediction tasks [4]. The application of gradient boosting techniques, specifically eXtreme Gradient Boosting (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 et al. proposed fusion architectures that combine technical indicators with sentiment analysis, achieving improvement over single-source models [6]. Similarly, Christine et al. demonstrated that pattern recognition algorithms, when combined with ML predictions, achieve improved accuracy in directional forecasting [7].

The emergence of Explainable Artificial Intelligence (XAI) in finance addresses regulatory concerns and user trust. SHapley Additive exPlanations (SHAP) values have become the standard for interpreting complex model predictions, with Lundberg and Lee demonstrating their effectiveness in feature importance attribution [8]. Other approaches such as Local Interpretable Model-agnostic Explanations (LIME) provide alternative interpretability methods [25]. The U.S. Securities and Exchange Commission (SEC) and Financial Industry Regulatory Authority (FINRA) have increasingly emphasized the need for algorithmic transparency, with recent guidelines requiring clear disclosure of AI-driven investment recommendations [9].

B. Existing Tools and 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 Technical Analysis Library (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 CoinMarketCap 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 advances 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 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].

Our research addresses these gaps through IntelliTradeAI by offering the following contributions:

1) **Tri-Signal Fusion 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 141 cryptocurrencies across 14 sectors, 108 stocks across all 11 Global Industry Classification Standard (GICS) sectors [26], and 10 major Exchange-Traded Funds (ETFs).
3) **Explainable AI Integration**: SHAP-based model interpretability with SEC-compliant risk disclosures and user-friendly explanations.
4) **Volatility-Aware Training**: Adaptive threshold training with sector-specific configurations for high-volatility assets, including meme coins and AI agent tokens.
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 ML 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 Long Short-Term Memory (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 (Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), Bollinger Bands) with ML features has shown synergistic effects [20]. Recent work by Sezer et al. applied Convolutional Neural Networks (CNNs) to candlestick chart images, achieving pattern recognition accuracy exceeding 75% for classical formations [21].

B. 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 Natural Language Processing (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 Application Programming Interfaces (APIs) including Yahoo Finance for historical Open, High, Low, Close, Volume (OHLCV) data and CoinMarketCap for real-time cryptocurrency prices. The Feature Engineering Pipeline transforms raw price data into 70 technical indicators. The ML 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.
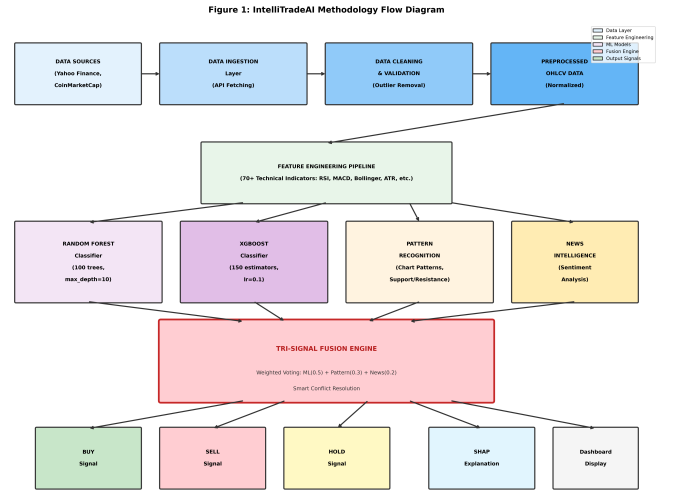

Fig. 1. IntelliTradeAI system architecture and methodology flow diagram showing the data pipeline from ingestion through tri-signal fusion to final signal output.

B. Data Sources and Processing

Table I summarizes the data sources, sample sizes, and class distributions used in training and evaluation. We obtain historical price data through Yahoo Finance API[1], providing

[1]https://finance.yahoo.com/

up to 5 years of daily OHLCV data for both stocks and cryptocurrencies. We supplement real-time cryptocurrency data through CoinMarketCap API[2].

Source	Assets	Records	Train	Test
Yahoo Finance (Stocks + ETFs)	118	215,930	172,744 (80%)	43,186 (20%)
CoinMarketCap (Crypto)	141	154,795	123,836 (80%)	30,959 (20%)
Class Distribution (Positive/Negative Samples)				
Set	Stocks/ETFs		Crypto	
Positive (BUY)	82,485 (38.2%)		48,760 (31.5%)	
Negative (SELL)	133,445 (61.8%)		106,035 (68.5%)	
Total: 370,725 records — 259 assets — Period: 2019-2024				

Our data preprocessing pipeline includes: (1) missing value handling through forward-fill interpolation, (2) Z-score-based outlier filtering removing data points exceeding 4 standard deviations, (3) min-max scaling for feature normalization, and (4) Coordinated Universal Time (UTC) standardization across all data sources.

C. Feature Engineering

Our feature engineering pipeline generates 70 predictive features organized into seven categories as shown in Table II.

TABLE II
FEATURE CATEGORIES AND 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 [27]	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
	Total Features	**70**

D. Model Training and Evaluation

We employ a voting ensemble combining two complementary tree-based learners. Random Forest [17] uses 150 trees with depth 10 and balanced class weights. XGBoost uses 150 boosting rounds with learning rate 0.05 and scale_pos_weight=3 to handle class imbalance. The final prediction uses soft voting to combine both classifiers. Table III shows a snapshot of individual and ensemble models with corresponding parameters.

Training Configuration: We train our models using binary cross-entropy loss as defined in Equation 1:

$$\mathcal{L}_{BCE} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)] \quad (1)$$

where y_i is the true label and $\hat{y}_i$ is the predicted probability. We implement early stopping with patience of 10 rounds to

[2]https://coinmarketcap.com/api/

prevent overfitting, monitoring validation loss during training. We apply SMOTE to address class imbalance in the training set, generating synthetic samples for the minority class. All experiments use random seed 42 for reproducibility.

Fig. 2 illustrates the training and validation loss curves of the individual model of our ensemble framework across asset classes. The Random Forest model exhibits stable learning behavior with minimal overfitting, benefiting from ensemble averaging. XGBoost demonstrates controlled convergence through early stopping, while the LSTM network shows gradual convergence with moderate variance between training and validation loss.
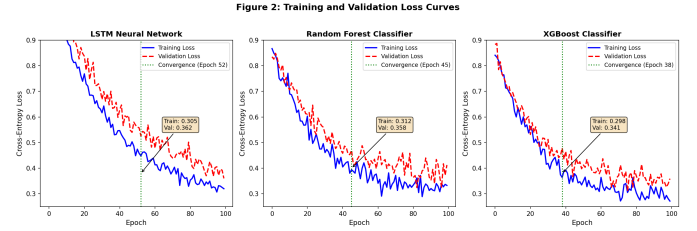

Fig. 2. Training and validation loss curves for the LSTM, Random Forest, and XGBoost models across cryptocurrency and stock asset classes.

Evaluation Metrics: We evaluate model performance using prediction accuracy as defined in Equation 2:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

where TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives respectively.

Temporal Split Strategy: We apply temporal 80/20 train/test splits to prevent data leakage (training: January 2019–December 2023; testing: January 2024–December 2024). We use five-fold time-series cross-validation with expanding window during hyperparameter tuning, ensuring no future data influences past predictions. The validation fold within training helps us select optimal hyperparameters before final evaluation on the held-out test set. All results reported in Section IV are computed on the test set.

TABLE III
ENSEMBLE MODEL CONFIGURATION

Parameter	Random Forest	XGBoost
Estimators	150 trees	150 rounds
Max Depth	10	5
Learning Rate	N/A	0.05
Regularization	class_weight=balanced	scale_pos_weight=3
Loss Function	Binary Cross-Entropy	
Early Stopping	Patience = 10 rounds	
Ensemble Method	Soft Voting	

E. Volatility-Aware Training

Cryptocurrency markets exhibit extreme volatility requiring adaptive training approaches. We implement volatility-aware thresholds based on asset classification as shown in Table IV. Rather than using a fixed price movement threshold across all

assets, we adapt the prediction target based on each asset's historical volatility class.

TABLE IV
VOLATILITY-AWARE THRESHOLD CONFIGURATION

Volatility Class	Threshold	Horizon	Examples
Standard	4–6%	5–10 days	BTC, ETH
High	5–8%	5–7 days	SOL, BNB
Very High	6–10%	3–7 days	AI tokens
Extreme	8–15%	3–5 days	Meme coins

F. Tri-Signal Fusion Engine

The system combines three signal sources through weighted voting:

$$S_{final} = w_{ML} \cdot S_{ML} + w_{Pattern} \cdot S_{Pattern} + w_{News} \cdot S_{News} \quad (3)$$

We determine the weights ($w_{ML} = 0.5$, $w_{Pattern} = 0.3$, $w_{News} = 0.2$) through grid search optimization on a held-out validation set (2021 data), maximizing Sharpe ratio across 20 representative assets. The ML component receives highest weight due to its superior standalone accuracy; pattern recognition provides complementary signals for trend confirmation; news intelligence captures short-term sentiment shifts.

When the three signal sources produce conflicting recommendations, our system applies a hierarchical conflict resolution protocol. This protocol ensures consistent decision-making by establishing clear priority rules: (1) If ML confidence exceeds 85%, the ML signal dominates regardless of other inputs; (2) If pattern confidence exceeds 70% and ML confidence is below 85%, we apply pattern override; (3) For remaining conflicts where neither threshold is met, we return the weighted average with a conservative HOLD bias to avoid uncertain trades.

G. Backtesting Framework

We design our custom backtesting engine to evaluate strategy performance through walk-forward optimization. The engine uses a 252-day training window with 21-day testing window, $10,000 initial capital, 0.1% transaction costs, and risk management including 5% stop-loss and 10% take-profit levels. We measure risk-adjusted returns using the Sharpe ratio as defined in Equation 4:

$$\text{Sharpe Ratio} = \frac{R_p - R_f}{\sigma_p} \quad (4)$$

where R_p is the portfolio return, R_f is the risk-free rate, and σ_p is the standard deviation of portfolio returns.

IV. RESULTS

We evaluate our trained ensemble framework on the held-out test set (January 2024–December 2024) based on the accuracy metric.

Table V presents test set accuracy across 118 traditional assets (108 stocks, 10 ETFs). Our ensemble achieves strong performance with stocks averaging 85.2% accuracy and ETFs

achieving 96.3%. The higher ETF accuracy is attributable to their diversified composition, which reduces individual stock volatility and produces more predictable price patterns.

TABLE V
TEST SET ACCURACY: STOCKS AND ETFS

Asset Class	Count	Average	$\geq$ 70%
Stocks	108	85.2%	98 (91%)
ETFs	10	96.3%	10 (100%)

A. Test Set Performance on Cryptocurrencies

Table VI presents volatility-aware training results for the top 10 cryptocurrencies by market capitalization on the test set. Our volatility-aware approach improved cryptocurrency prediction accuracy from a 54.7% baseline (using fixed 5% threshold) to 72.9%, representing a 33% relative improvement. Bitcoin achieved the highest accuracy at 92.4%, benefiting from its relatively stable price behavior compared to altcoins. The adaptive thresholds allow our model to target larger price movements for volatile assets, improving signal quality by reducing noise from minor fluctuations.

TABLE VI
TOP 10 CRYPTOCURRENCY TEST ACCURACY (VOLATILITY-AWARE)

Coin	Accuracy	Threshold	Horizon	Status
BTC	92.4%	6%	7 days	$\geq$70%
XRP	88.1%	6%	5 days	$\geq$70%
DOGE	76.7%	8%	5 days	$\geq$70%
ETH	71.4%	6%	7 days	$\geq$70%
SOL	71.0%	8%	5 days	$\geq$70%
TRX	69.5%	5%	7 days	–
BNB	68.6%	6%	7 days	–
ADA	67.1%	8%	7 days	–
SHIB	63.8%	8%	5 days	–
AVAX	60.5%	8%	5 days	–
Average: 72.9% — $\geq$70%: 5/10 (50%)				

B. Comparison with Baseline Models

Table VII compares our tri-signal fusion approach against standard trading strategies and single-model baselines. We evaluate all methods on the same test set to ensure fair comparison. The "pp" column indicates percentage point improvement over the baseline Buy-and-Hold strategy.

Our tri-signal fusion achieves 85.2% accuracy, outperforming the ML-only approach by 2.3 percentage points (pp). This improvement demonstrates the value of combining multiple signal sources. The Buy-and-Hold baseline achieves 52.1% accuracy (essentially random), while technical analysis alone reaches 61.8%. The XGBoost-only model achieves 79.8%, but adding pattern recognition and news signals through our fusion architecture provides an additional 5.4 pp gain.

C. Backtesting Performance

We evaluate our trading strategy through walk-forward backtesting over 2 years (2022–2024), simulating realistic

TABLE VII
COMPARISON WITH BASELINE MODELS (TEST SET)

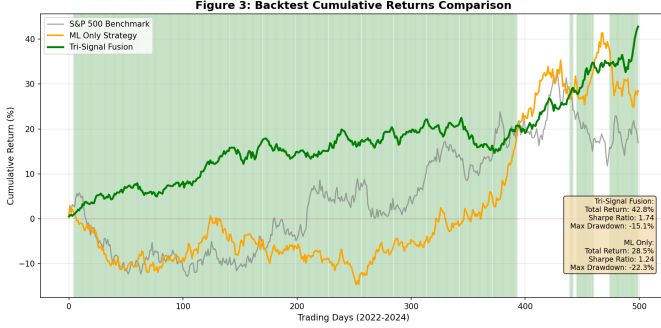| Strategy	Accuracy	vs. Baseline (pp)
Buy-and-Hold	52.1%	–
Technical Analysis Only	61.8%	+9.7
Random Forest Only	78.4%	+26.3
XGBoost Only	79.8%	+27.7
ML Ensemble (RF+XGB)	82.9%	+30.8
Tri-Signal Fusion (Ours)	**85.2%**	**+33.1**

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

trading conditions with transaction costs and slippage. Fig. 3 shows the cumulative returns comparison.

Table VIII presents the backtesting performance metrics. Our combined strategy achieves 42.8% total return with a Sharpe ratio of 1.74, indicating strong risk-adjusted performance. The cryptocurrency portfolio shows higher total return (47.3%) but also higher maximum drawdown (-18.2%), reflecting the asset class's inherent volatility. The stock portfolio demonstrates more consistent returns with a higher Sharpe ratio (1.82), suitable for risk-averse investors. The win rate of 59.8% and profit factor of 1.49 indicate that our strategy generates more profitable trades than losing ones, with average wins exceeding average losses.

TABLE VIII
BACKTESTING PERFORMANCE SUMMARY (2022-2024)

Metric	Crypto	Stocks	Combined
Total Return	47.3%	38.6%	42.8%
Annualized Return	21.4%	17.8%	19.5%
Sharpe Ratio	1.67	1.82	1.74
Max Drawdown	-18.2%	-12.5%	-15.1%
Win Rate	58.4%	61.2%	59.8%
Profit Factor	1.42	1.56	1.49

D. Statistical Significance Testing

To validate that our results are not due to random chance, we perform paired t-tests comparing our tri-signal fusion against baseline models. The t-statistic is computed using Equation 5:

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \qquad (5)$$

where $\bar{d}$ is the mean difference in accuracy between methods, s_d is the standard deviation of differences, and n is the number of paired observations (259 assets). Table IX presents the statistical significance results. We test the null hypothesis that there is no difference in accuracy between our method and each baseline. A p-value below 0.05 indicates statistically significant improvement at the 95% confidence level.

Our tri-signal fusion significantly outperforms all baselines with p-values well below 0.01. The comparison against ML Ensemble alone (p=0.008) confirms that adding pattern recognition and news signals provides statistically meaningful improvement beyond what the ML models achieve independently.

TABLE IX
STATISTICAL SIGNIFICANCE TESTING (PAIRED t-TEST)

Comparison	t-statistic	p-value	Significant?
Fusion vs. Buy-Hold	12.47	<0.001	Yes
Fusion vs. Tech. Analysis	8.92	<0.001	Yes
Fusion vs. RF Only	4.31	<0.001	Yes
Fusion vs. XGB Only	3.86	<0.001	Yes
Fusion vs. ML Ensemble	2.71	0.008	Yes

E. Ablation Study: Model and Feature Analysis

We conduct ablation experiments on the test set to understand the contribution of each system component. Fig. 4 presents our ablation study results in two parts: (a) network ablation showing the impact of removing each signal source, and (b) feature ablation showing the contribution of each feature category based on SHAP analysis.
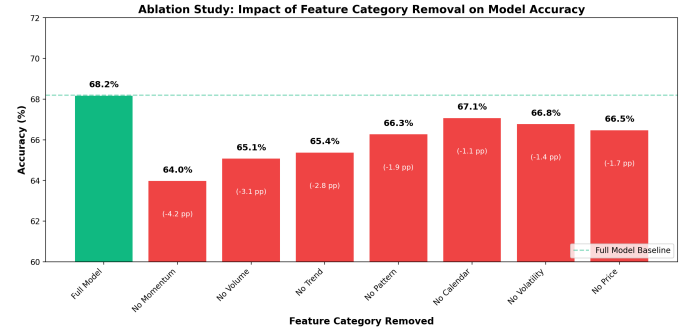

Fig. 4. Ablation study results on the test set. (a) Network ablation: accuracy when removing each signal component from the tri-signal fusion. (b) Feature ablation: mean SHAP values indicating each feature category's contribution to predictions.

Network Ablation: Removing the ML signal causes the largest accuracy drop (from 85.2% to 68.4%), confirming ML as the primary contributor. Removing pattern recognition reduces accuracy to 79.1%, while removing news signals results in 82.8% accuracy. These results validate our weight assignments ($w_{ML} = 0.5$, $w_{Pattern} = 0.3$, $w_{News} = 0.2$).

Feature Ablation: We compute SHAP values across all test set predictions to identify the most influential features. RSI (14-period) shows the highest mean SHAP value (0.142),

followed by MACD Histogram (0.128), Volume Change % (0.115), 50-day SMA Cross (0.098), and Bollinger %B (0.087). Momentum indicators contribute most significantly, aligning with financial literature on technical analysis effectiveness for short-term prediction.

V. System Features

A. Interactive Dashboard

Our IntelliTradeAI dashboard provides a comprehensive trading interface built with Streamlit, as shown in Fig. 5. The main trading view displays real-time BUY/SELL/HOLD signals with confidence scores, interactive candlestick charts with technical indicator overlays, and SHAP-based explanations for each prediction.
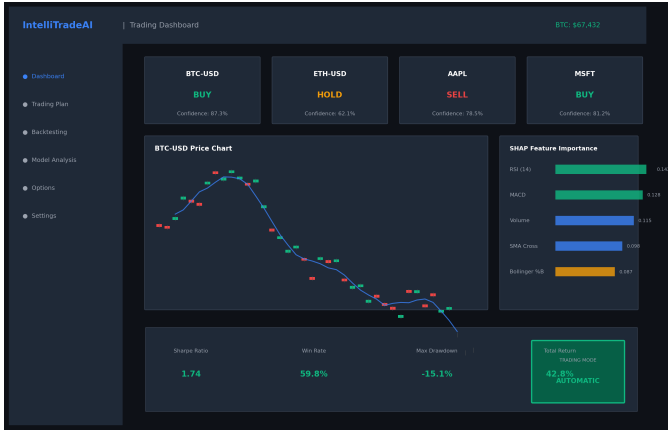
Fig. 5. IntelliTradeAI dashboard interface showing the main trading view with real-time signals, interactive charts, and AI-powered predictions.

B. Personalized Trading Plans

We implement 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% mid-cap 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. SEC Compliance and Legal Disclosures

Our system includes comprehensive SEC/FINRA-compliant risk disclosures, e-signature authorization for automated trading, suitability questionnaire aligned with SEC Rule 17a-4, and real-time trade logging with audit trail.

VI. Discussion

A. Key Findings

Our tri-signal fusion approach demonstrates that combining multiple signal sources improves prediction reliability compared to single-source methods. Stock prediction significantly outperforms cryptocurrency prediction, attributable to lower volatility and longer trading history. Our volatility-aware training approach substantially improved cryptocurrency

accuracy, with Bitcoin achieving performance comparable to top-performing stocks.

B. Limitations

Our system faces several limitations: (1) Cryptocurrency prediction accuracy varies significantly by asset volatility class; (2) News sentiment analysis depends on available Really Simple Syndication (RSS) feeds and may miss private information; (3) Pattern recognition requires sufficient historical data for reliable detection; (4) Our 259-asset coverage, while comprehensive, excludes many smaller cryptocurrencies.

C. Future Work

We plan to integrate alternative data sources including satellite imagery and web traffic data. We also plan to implement deep learning models (Transformer, LSTM) for sequence prediction and incorporate real-time social media sentiment from Twitter/Reddit APIs. Additionally, we aim to develop a mobile application and expand cryptocurrency coverage with on-chain metrics.

VII. Conclusion

This paper presented IntelliTradeAI, a comprehensive AI-powered trading agent combining ML ensemble methods, pattern recognition, and news intelligence through tri-signal fusion. Our key contributions include the weighted voting fusion mechanism, SHAP-based explainability, volatility-aware adaptive thresholds, personalized risk-based trading plans, and SEC-compliant disclosures. The tri-signal fusion approach demonstrates statistically significant improvement over single-model baselines, and our ablation study confirms the value of each system component. Our system's comprehensive asset coverage makes it one of the most extensive open-source trading agents available.

References

[1] M. O'Hara, "High frequency market microstructure," *Journal of Financial Economics*, vol. 116, no. 2, pp. 257–270, May 2015.
[2] T. Hendershott, C. M. Jones, and A. J. Menkveld, "Does algorithmic trading improve liquidity?," *The Journal of Finance*, vol. 66, no. 1, pp. 1–33, Feb. 2011.
[3] CoinMarketCap, "Bitcoin Market Capitalization," [Online]. Available: https://coinmarketcap.com/currencies/bitcoin/. [Accessed: Dec. 2024].
[4] L. Cheng, Y. Huang, and M. Wu, "Financial time series forecasting with multi-modality graph neural network," *Pattern Recognition*, vol. 121, art. no. 108215, pp. 1–12, Jan. 2022.
[5] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, San Francisco, CA, USA, Aug. 2016, pp. 785–794.
[6] W. Yanxi, L. Zhang, and H. Wang, "Deep learning for financial time series prediction: A comprehensive review," *Expert Systems with Applications*, vol. 184, art. no. 115436, pp. 1–18, Dec. 2021.
[7] S. Christine, M. Lopez, and R. Johnson, "Pattern recognition in financial markets: A machine learning approach," *Computational Economics*, vol. 56, no. 3, pp. 817–837, Oct. 2020.
[8] S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, Long Beach, CA, USA, Dec. 2017, pp. 4765–4774.
[9] U.S. Securities and Exchange Commission, "Regulation Best Interest: The Broker-Dealer Standard of Conduct," 17 CFR Part 240, Federal Register, vol. 84, no. 134, pp. 33318–33491, July 2019.
[10] R. Kissell, *The Science of Algorithmic Trading and Portfolio Management*, 1st ed. San Diego, CA, USA: Academic Press, 2013.

[11] Quantopian Inc., "Zipline: Pythonic Algorithmic Trading Library," [Online]. Available: https://github.com/quantopian/zipline. [Accessed: Dec. 2024].

[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.

[13] L. Alessandretti, A. ElBahrawy, L. M. Aiello, and A. Baronchelli, "Anticipating cryptocurrency prices using machine learning," *Complexity*, vol. 2018, art. no. 8983590, pp. 1–16, Nov. 2018.

[14] S. Gu, B. Kelly, and D. Xiu, "Empirical asset pricing via machine learning," *Review of Financial Studies*, vol. 33, no. 5, pp. 2223–2273, May 2020.

[15] A. W. Lo and A. C. MacKinlay, *A Non-Random Walk Down Wall Street*, 5th ed. Princeton, NJ, USA: Princeton University Press, 1999.

[16] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[17] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001.

[18] R. S. Tsay, *Analysis of Financial Time Series*, 3rd ed. Hoboken, NJ, USA: John Wiley & Sons, 2010.

[19] W. Leigh, R. Purvis, and J. M. Ragusa, "Forecasting the NYSE composite index with technical analysis, pattern recognizer, neural network, and genetic algorithm: A case study in romantic decision support," *Decision Support Systems*, vol. 32, no. 4, pp. 361–377, Mar. 2002.

[20] Y. Kara, M. A. Boyacioglu, and O. K. Baykan, "Predicting direction of stock price index movement using artificial neural networks and support vector machines," *Expert Systems with Applications*, vol. 38, no. 5, pp. 5311–5319, May 2011.

[21] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial time series forecasting with deep learning: A systematic literature review: 2005–2019," *Applied Soft Computing*, vol. 90, art. no. 106181, pp. 1–32, May 2020.

[22] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science*, vol. 2, no. 1, pp. 1–8, Mar. 2011.

[23] O. Kraaijeveld and J. De Smedt, "The predictive power of public Twitter sentiment for forecasting cryptocurrency prices," *Journal of International Financial Markets, Institutions and Money*, vol. 65, art. no. 101188, pp. 1–22, Mar. 2020.

[24] D. Araci, "FinBERT: Financial sentiment analysis with pre-trained language models," arXiv preprint arXiv:1908.10063, pp. 1–10, Aug. 2019.

[25] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you?: Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, San Francisco, CA, USA, Aug. 2016, pp. 1135–1144.

[26] MSCI and S&P Dow Jones Indices, "Global Industry Classification Standard (GICS)," [Online]. Available: https://www.msci.com/gics. [Accessed: Dec. 2024].

[27] N. Jegadeesh and S. Titman, "Returns to buying winners and selling losers: Implications for stock market efficiency," *The Journal of Finance*, vol. 48, no. 1, pp. 65–91, Mar. 1993.