Finance Research Letters 58 (2023) 104285

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

Stop-loss adjusted labels for machine learning-based trading of risky assets

Yoontae Hwang [a], Junpyo Park [a], Yongjae Lee [a,*], Dong-Young Lim [a,b,*]

[a] Department of Industrial Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea
[b] Artificial Intelligence Graduate School, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea

ARTICLE INFO

Keywords:
Stop-loss trading
Asset price prediction
Cryptocurrency
Machine learning

ABSTRACT

Since the rise of ML/AI, many researchers and practitioners have been trying to predict future stock price movements. In actual implementations, however, stop-loss is widely adopted to manage risks, which sells an asset if its price goes below a predetermined level. Hence, some buy signals from prediction models could be wasted if stop-loss is triggered. In this study, we propose a stop-loss adjusted labeling scheme to reduce the discrepancy between prediction and decision making. It can be easily incorporated to any ML/AI prediction models. Experimental results on U. S. futures and cryptocurrencies show that this simple tweak significantly reduces risk.

1. Introduction

In today's financial markets, traders face increased volatility, uncertainty, and risk (Liu and Serletis, 2019). Effective risk management strategies are essential for traders to safely achieve their investment objectives. One such strategy is stop-loss trading, which allows traders to limit their potential losses by setting a predetermined price at which to sell a security (Han et al., 2016). Many other strategies often incorporate stop-loss orders to prevent unlimited losses. For example, mean-reversion (Jegadeesh, 1990, 1991; Li et al., 2017) and grid trading (Rundo et al., 2019; Yeh et al., 2022). Although stop-loss trading strategies are widely adopted by traders and industrial professionals, there have been relatively few empirical studies on the effectiveness of the stop-loss mechanism. (Lei and Li, 2009; Lo and Remorov, 2017; Dai et al., 2021).

Previous researchers have attempted to predict future asset prices using ML/AI models[3], but they have only focused on predicting the asset price at $t + \Delta$, ignoring information between $(t, t + \Delta)$ (Patel et al., 2015; Gonçalves et al., 2019). However, some buy signals may be wasted for stop-loss traders because an asset should be sold if its price goes below the stop-loss price during $(t, t + \Delta)$ regardless of its price at $t + \Delta$. Hence, there is a discrepancy between prediction and decision-making.

To make prediction and decision-making more consistent, we propose a stop-loss adjusted labeling scheme, which can be easily incorporated into various ML/AI prediction models. Experiments using various ML/AI models on U.S. futures and cryptocurrencies suggest that this simple modification can significantly reduce risk.

* Corresponding authors at: Department of Industrial Engineering, Ulsan National Institute of Science and Technology (UNIST), 50 UNIST-gil, Ulju-gun, Ulsan 44919, Republic of Korea.

E-mail addresses: yongjaelee@unist.ac.kr (Y. Lee), dlim@unist.ac.kr (D.-Y. Lim).

https://doi.org/10.1016/j.frl.2023.104285
Received 8 June 2023; Received in revised form 18 July 2023; Accepted 27 July 2023
Available online 31 July 2023
1544-6123/

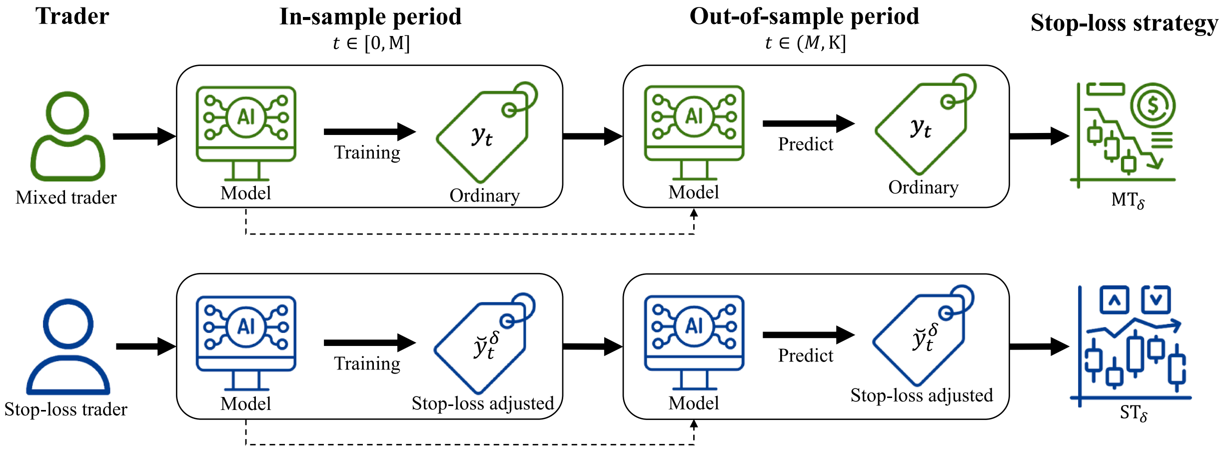

Fig. 1. Investment process of a mixed trader (**MT**$_\delta$) and a stop-loss trader(**ST**$_\delta$).

2. Methodology

2.1. Target labels

Let $p_t \in R_+$ be the price of an asset at time t. Most of the previous studies on predicting directional movements of asset prices (Gu et al., 2020; Sirignano, 2019; Sirignano and Cont, 2019; Ju et al., 2019; Ghosh et al., 2022; Sun et al., 2020) use the following (ordinary) label: for a fixed time step $\Delta > 0$,

$$y_t = \begin{cases} 1, & \text{if } \frac{p_{t+\Delta}}{p_t} > 1. \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Now let us propose a ***stop-loss adjusted label***. If we let $\delta \in [0, 1)$ be the maximum tolerance level for stop-loss trading and m_t^Δ be the lowest value of the asset price during $[t, t+\Delta]$, i.e., $m_t^\Delta := \min\{p_s \mid s \in [t, t+\Delta]\}$. then, a stop-loss adjusted label $\breve{y}_t^\delta$ can be defined as

$$\breve{y}_t^\delta = \begin{cases} 1, & \text{if } \left\{\frac{p_{t+\Delta}}{p_t} > 1\right\} \cap \left\{\frac{m_t^\Delta}{p_t} \geq (1-\delta)\right\} \\ 0, & \text{otherwise.} \end{cases} . \tag{2}$$

Note that $y_t = \breve{y}_t^\delta$ when $\delta = 1$. The value of δ should be set depending on the risk preference of a trader. This paper focuses on a stop-loss adjusted labeling scheme for buy signals, but a similar approach can be applied to short positions.

2.2. Trader types

We assume that assets can be traded at discrete time points $\tau = \{t_0, t_1, \ldots, t_K\}$ with $t_i = t + i\Delta$ for $i \in \{0, \ldots, K\}$. We consider traders who employ a stop-loss trading strategy (with threshold δ). It needs a trading signal $\tilde{y}_t$, which can be based on the ordinary labeling y_t or the stop-loss labeling $\breve{y}_t^\delta$. The signal-based stop-loss strategy can be described as follows:

If $\tilde{y}_t = 1$: Buys the asset at time t.

During time $s \in (t, t+\Delta)$, as soon as $p_s/p_t < 1-\delta$, sells the asset.
Otherwise, sells the asset at time $t+\Delta$.

If $\tilde{y}_t = 0$: Do nothing.

In this study, we compare the performance of two different types of traders: mixed trader and stop-loss trader. While both traders employ the stop-loss trading strategy, the mixed trader (MT_δ) uses ordinary label y_t and the stop-loss trader (ST_δ) uses stop-loss adjusted label $\breve{y}_t^\delta$. Fig. 1 illustrates the investment process of the two traders.

Table 1
Descriptive statistics of selected assets: US futures and cryptocurrencies.

Panel A. U.S. futures												
Asset	**Mean**		**Std dev**		**SR**		**MDD**		**VaR 95%**		**CVaR 95%**	
	IS	**OOS**	**IS**	**OOS**	**IS**	**OOS**	**IS**	**OOS**	**IS**	**OOS**	**IS**	**OOS**
YM	0.008	0.001	0.722	0.522	0.010	0.001	0.765	0.191	0.901	0.754	1.761	1.414
NQ	0.011	-0.005	0.624	0.786	0.018	-0.006	0.580	0.351	0.880	1.161	1.587	2.081
ES	0.007	0.001	0.682	0.597	0.011	0.001	0.741	0.246	0.870	0.870	1.693	1.588
PA	0.011	-0.011	0.909	1.566	0.012	-0.007	0.654	0.499	1.275	2.243	2.193	3.979
EW	0.010	-0.001	0.836	0.726	0.013	-0.001	0.683	0.261	0.961	1.104	1.997	1.925
UB	0.005	-0.018	0.429	0.490	0.011	-0.036	0.268	0.296	0.616	0.810	1.003	1.227
CL	-0.038	0.035	5.685	1.475	-0.007	0.024	1.021	0.303	1.412	2.254	4.579	3.881
VX	-0.012	-0.045	0.686	1.926	-0.018	-0.023	0.730	0.621	0.810	2.860	1.523	4.333
RTY	0.009	-0.009	0.849	0.761	0.010	-0.012	0.676	0.342	1.106	1.218	2.109	1.993
Panel B. Cryptocurrency												
Asset	**Mean**		**Std dev**		**SR**		**MDD**		**VaR 95%**		**CVaR 95%**	
	IS	**OOS**	**IS**	**OOS**	**IS**	**OOS**	**IS**	**OOS**	**IS**	**OOS**	**IS**	**OOS**
BTC	0.027	-0.022	1.826	1.466	0.015	-0.015	0.898	0.770	2.714	2.352	4.566	3.488
ETH	0.024	-0.015	2.237	1.886	0.011	-0.008	0.974	0.849	3.359	3.110	5.596	4.459
XRP	-0.001	-0.031	2.384	1.969	0.000	-0.016	0.930	0.845	3.291	3.112	5.835	4.598
SOL	0.125	-0.004	4.093	2.763	0.031	-0.001	0.849	0.932	5.926	4.247	9.016	6.202
BNB	0.066	-0.007	2.628	1.775	0.025	-0.004	0.917	0.768	3.719	2.864	6.276	4.250
ADA	0.023	-0.045	2.409	2.172	0.010	-0.021	0.979	0.914	3.620	3.346	5.869	4.991
DOGE	0.104	-0.057	3.529	2.428	0.029	-0.024	0.784	0.913	3.281	3.721	7.047	5.485

This table reports the descriptive statistics of returns of U.S. futures and cryptocurrencies, including their mean, standard deviation (Std dev), Sharpe ratio (SR), maximum drawdown (MDD), value at risk (VaR) at 95%, and conditional value at risk (CVaR) at 95%. IS and OOS refer to in-sample and out-of-sample period.

Table 2
Input features.

Input features	Description
z_{open}	$\text{open}_t/\text{close}_t - 1$
z_{high}	$\text{high}_t/\text{close}_t - 1$
z_{low}	$\text{low}_t/\text{close}_t - 1$
z_{close}	$\text{close}_t/\text{close}_{t-1} - 1$
z_k	$\frac{\sum_{i=0}^{k}\text{close}_{t-i}}{k\cdot\text{close}_t} - 1$ with $k = \{5, 10, 15, 20, 25, 30\}$
Time variables	Weekdays, months
Technical indicators	MACD (of price and volume), RSI, OBV

3. Experiment

3.1. Experiment settings

In this section, we present details on the datasets, preprocessing, classification settings, and hyperparameter tuning employed in our experiments.

3.1.1. Datasets

We use two distinct asset classes: U.S. futures and cryptocurrencies. For each asset class, we choose individual assets with large volumes and market capitalizations. The description for all asset tickers can be found in appendix A.

Table 1 provides descriptive statistics of selected assets. For U.S. futures, the in-sample period ranges from May 2, 2006 to June 20, 2021, and the out-of-sample period ranges from June 21, 2021 to August 26, 2022. For cryptocurrencies, the in-sample period ranges from August 22, 2017 to June 20, 2021, and the out-of-sample period is identical to that of U.S. futures. The in-sample period for our dataset was carefully chosen to encompass both bear and bull markets (namely, the 2008 global financial crisis and the subsequent record bull run), ensuring a fair comparison between ordinary and stop-loss adjusted labels. This selection helps minimize biases that could arise if the training period were dominated by one type of market condition. The out-of-sample period includes the COVID-19 period to evaluate the effectiveness of stop-loss strategies during extraordinary market conditions.

3.1.2. Preprocessing

Table 2 provides detailed information on the input features: OHLC (Open, High, Low, Close), time variables, and technical indicators. The OHLC values are normalized with respect to the closing price, as shown in Table 2. In addition, we incorporate z_k to reflect both short- and long-term price trends (Feng et al., 2019; Yoo et al., 2021). We select three popular technical indicators from

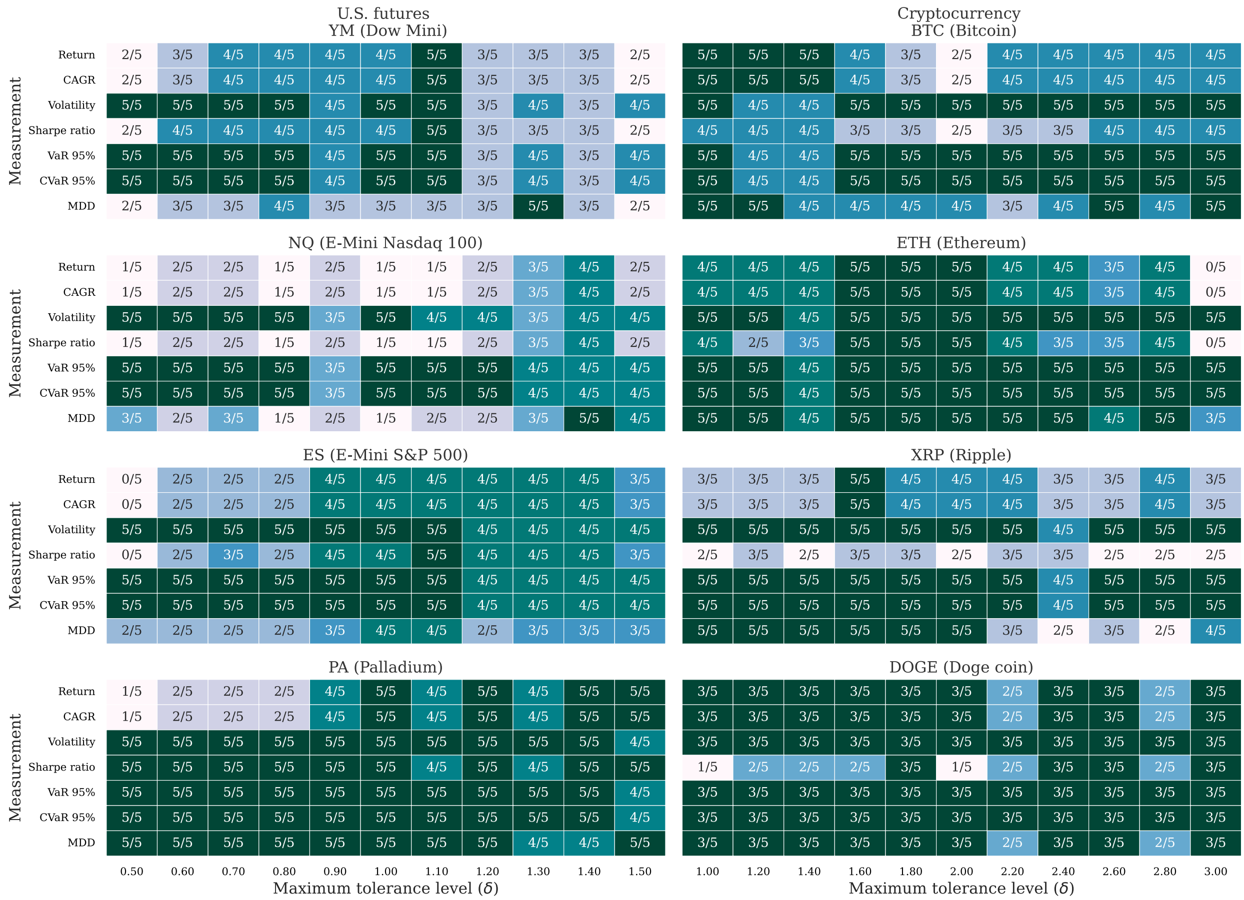

Fig. 2. Win-Matrices of $\mathbf{ST}_\delta$ and $\mathbf{MT}_\delta$ on four U.S. futures and four cryptocurrencies.

Santos and Torrent (2022): Moving Average Convergence/Divergence (MACD), Relative Strength Index (RSI), and On-Balance-Volume (OBV). MACD is used for both price and volume. Note that these conventional (OHLC and time) and technical (MACD, RSI, and OBV) features are widely used for identifying patterns in asset price movements (e.g., Nelsen et al., 2017; Gao and Chai, 2018). All features are calculated at 4-hour intervals, i.e., $\Delta = 4$-hour.

3.1.3. Classification setting

Consider time series data $\mathscr{X}_t = \{X_{t-i}\}_{i=0}^{M}$ with $X_{t-i} \in R^C$, where M is the lookback period, X_t is the vector of feature variables at time t, and C is the total number of features. The objective is to predict the label y_t (or $\check{y}_t^\delta$) given $\mathscr{X}_t$.

3.1.4. Prediction models

For our experiment, we use five popular classification models: multi-layer perceptron (MLP), XGBoost (XG), random forest (RF), CatBoost (CB), and k-nearest neighbors classifier (KNN). Note that MLP is a deep learning model, and the other four are classical machine learning models.

3.1.5. Hyperparameter tuning

The hyperparameters of prediction models are optimized based on Akiba et al. (2019). For the four classical machine learning models, we used a tree-structured Parzen estimator search on the cross-validation split. For the MLP model, we find the optimal hyperparameters by a random grid search.

3.2. Type of investors and their profits

In Section 2, we define two types of investors: mixed trader (MT_δ) and stop-loss trader (ST_δ). Assume that the two traders MT_δ and ST_δ have the trained models to predict ordinary label y_t and stop-loss adjusted label $\check{y}_t^\delta$, respectively. Here, we assume no transaction cost and slippage for simplicity. Then, the profit of MT_δ over the period $[t, t+\Delta]$ can be written as follows:

Fig. 3. Showcase examples.

$$\mathrm{PE}_t^{\mathrm{MT}_\delta} = 1_{y_t=1}\left(1_{m_t^\Delta \geq p_t(1-\delta)}\left(\frac{p_{t+\Delta}}{p_t}-1\right) - 1_{m_t^\Delta < p_t(1-\delta)}\delta\right). \tag{3}$$

Any transaction occurs only when the predicted label y_t is 1. Over the interval $[t, t+\Delta]$, if the minimum price m_t^Δ is greater or equal to the stop-loss price $p_t(1-\delta)$, stop-loss is not triggered, and thus the profit is represented by the relative price change $p_{t+\Delta}/p_t - 1$. Otherwise, if m_t^Δ is less than $p_t(1-\delta)$, stop-loss is triggered. Then, the profit is $-\delta$, which indicates a loss of δ. Similarly, the one-period profit of ST_δ is as follows:

$$\mathrm{PE}_t^{\mathrm{ST}_\delta} = 1_{\bar{y}_t^{\delta}=1}\left(1_{m_t^\Delta \geq p_t(1-\delta)}\left(\frac{p_{t+\Delta}}{p_t}-1\right) - 1_{m_t^\Delta < p_t(1-\delta)}\delta\right). \tag{4}$$

Here, $1_{(\cdot)}$ is an indicator function that becomes 1 if $(\cdot)$ is true and 0 otherwise. Recall that $m_t^\Delta := \min\{p_s \mid s \in [t,\ t+\Delta]\}$.

Throughout the experiments, we compare the accumulated profits of MT_δ and ST_δ during the out-of-sample period. We denote them by $\mathrm{PE}_\delta^{\mathrm{MT}}$ and $\mathrm{PE}_\delta^{\mathrm{ST}}$, respectively, which are summations of one-period profits of MT_δ and ST_δ over the out-of-sample period.

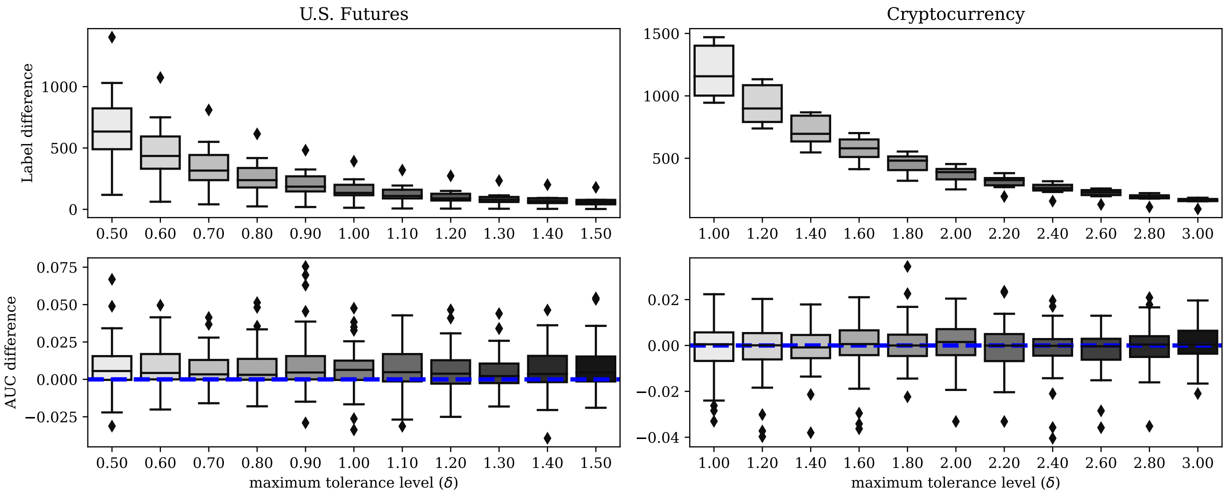

Fig. 4. Performance comparison of classification models with and without stop-loss adjustment Across Assets.

3.3. Experiment results

3.3.1. Portfolio performances

In this section, we analyze the portfolio performance of MT_δ and ST_δ during the out-of-sample period using several measures: annualized return, compound annual growth rate (CAGR), volatility, Sharpe ratio, maximum drawdown (MDD), Value at Risk at a 95% confidence level (VaR 95%), and Conditional Value at Risk at a 95% confidence level (CVaR 95%).

To effectively represent our extensive experiment results, we define a *win-matrix*. The win matrix is a 7-by-11 matrix, where the rows represent 7 different performance measures and the columns represent 11 different maximum tolerance levels for the stop-loss trading (δ). Each entry $W_{i,j}$ of the matrix is calculated as follows:

$$W_{i,j} = \frac{1}{5}\sum_{m=1}^{5} 1_{\left(PE_{i,j,m}^{ST_\delta} > PE_{i,j,m}^{MT_\delta}\right)}. \tag{5}$$

Here, $PE_{i,j,m}^{ST_\delta}$ and $PE_{i,j,m}^{MT_\delta}$ represent the performance of ST_δ and MT_δ with performance measure *i*, tolorence level *j*, and prediction model *m*. Therefore, each entry $W_{i,j}$ represents the winning rate of ST_δ over MT_δ given performance measure *i* and tolerance level *j* when tested with five different models (MLP, XG, RF, CB, and KNN).

Fig. 2 represents the win-matrix results for four representative U.S. futures and four representative cryptocurrencies.[1] First of all, we can see that ST_δ generally outperforms MT_δ in terms of risk measures such as volatility, VaR, CVaR, and MDD. That is, we can significantly reduce risk by simply adjusting labels to consider stop-loss. Furthermore, we observe that the stop-loss adjusted label is particularly effective for cryptocurrencies compared to U.S. futures. For cryptocurrencies, ST_δ outperforms MT_δ across most performance measures and tolerance levels. For U.S. futures, however, ST_δ outperforms MT_δ in terms of return-related measures such as annualized return, CAGR, and Sharpe ratio, but only within some specific range of δ. Due to the higher volatility of the cryptocurrency market, stop-loss is more likely to be triggered, and the average number of correct signals by ML models is also higher in cryptocurrencies than futures (Appendix A, Table A.2). As a result, the experiment results indicate that the stop-loss adjusted labeling scheme is more effective for cryptocurrencies compared to U.S. futures.

There have been some studies suggesting that stop-loss strategies primarily reduce risk rather than improve returns (Lei and Li, 2009; Kaminski and Lo, 2014; Lo and Remorov, 2017). Now our experiments show that a labeling scheme reflecting stop-loss can help investors to truly achieve the benefits of stop-loss strategies.

3.3.2. Showcase examples

We present some showcase examples of how stop-loss adjusted labels can reduce risk. The main purpose of stop-loss adjusted labels is to prevent buying assets that are expected to hit the stop-loss price. Hence, there could be some periods that ordinary labels give signals to buy assets, while stop-loss adjusted labels do not generate buy signals. Then, ST_δ can significantly reduce losses compared to MT_δ.

Fig. 3 shows some notable examples of such desirable cases. We can see that during the gray-shaded periods, MT_δ loses a lot while ST_δ successfully avoids losses.

[1] More detailed results for all assets (9 U.S. futures and 7 cryptocurrencies) are provided in Appendix B. The results are similar.

Table 3
Shapely value of input variables of XGBoost with different values of stop-loss threshold (δ).

δ	z_{open}	z_{high}	z_{low}	z_{close}	z_{d5}	z_{d10}	z_{d15}	z_{d20}	z_{d25}	z_{d30}	*RSI*	MACD_P	MACD_V	OBV
Panel A. U.S. futures														
0.50	0.266	**0.652**	0.292	0.179	0.268	0.328	**0.667**	0.378	0.473	0.242	0.490	0.264	0.369	0.576
0.60	0.238	**0.561**	0.230	0.202	0.313	0.362	**0.651**	0.519	0.486	0.315	0.504	0.276	0.311	0.418
0.70	0.191	0.430	0.159	0.232	0.344	0.422	**0.701**	0.426	0.480	0.327	**0.564**	0.293	0.248	0.278
0.80	0.228	0.422	0.147	0.277	0.310	0.398	**0.690**	0.463	0.409	0.342	**0.597**	0.294	0.224	0.207
0.90	0.244	0.377	0.141	0.322	0.393	0.479	**0.715**	0.511	0.464	0.359	**0.559**	0.313	0.238	0.170
1.00	0.232	0.366	0.132	0.316	0.382	0.550	**0.683**	0.540	0.565	0.339	**0.615**	0.313	0.221	0.105
1.10	0.205	0.315	0.113	0.303	0.338	0.495	**0.618**	0.468	0.487	0.325	**0.550**	0.289	0.222	0.067
1.20	0.228	0.293	0.103	0.297	0.332	0.509	**0.583**	0.411	0.487	0.339	**0.605**	0.299	0.214	0.050
1.30	0.275	0.282	0.091	0.312	0.336	0.537	**0.613**	0.459	0.506	0.321	**0.658**	0.294	0.216	0.040
1.40	0.239	0.248	0.081	0.300	0.341	0.502	**0.603**	0.421	0.597	0.368	**0.627**	0.316	0.193	0.020
1.50	0.308	0.260	0.073	0.285	0.348	0.548	**0.651**	0.366	0.597	0.364	**0.666**	0.345	0.216	0.013
Panel B. Cryptocurrency														
δ	z_{open}	z_{high}	z_{low}	z_{close}	z_{d5}	z_{d10}	z_{d15}	z_{d20}	z_{d25}	z_{d30}	*RSI*	MACD_P	MACD_V	OBV
1.00	0.196	0.203	0.235	**0.797**	0.172	0.336	0.044	0.075	0.047	0.066	0.059	0.127	0.152	**0.645**
1.20	0.204	0.218	0.231	**0.787**	0.172	0.388	0.052	0.070	0.058	0.061	0.086	0.121	0.172	**0.556**
1.40	0.172	0.208	0.231	**0.794**	0.166	0.365	0.038	0.059	0.030	0.064	0.065	0.113	0.163	**0.491**
1.60	0.185	0.217	0.224	**0.794**	0.169	0.372	0.030	0.032	0.023	0.057	0.058	0.110	0.146	**0.446**
1.80	0.197	0.233	0.207	**0.833**	0.159	0.382	0.041	0.065	0.019	0.065	0.057	0.122	0.124	**0.413**
2.00	0.199	0.229	0.199	**0.810**	0.155	0.401	0.067	0.074	0.019	0.063	0.057	0.137	0.133	**0.405**
2.20	0.212	0.265	0.220	**0.872**	0.152	0.356	0.063	0.071	0.025	0.082	0.058	0.148	0.121	**0.382**
2.40	0.179	0.254	0.194	**0.858**	0.134	**0.414**	0.067	0.041	0.022	0.065	0.041	0.153	0.100	0.373
2.60	0.198	0.269	0.211	**0.873**	0.176	**0.385**	0.078	0.054	0.024	0.056	0.071	0.142	0.123	0.355
2.80	0.203	0.267	0.206	**0.955**	0.163	**0.400**	0.101	0.064	0.030	0.047	0.079	0.164	0.114	0.358
3.00	0.230	0.275	0.241	**0.965**	0.175	**0.387**	0.094	0.085	0.042	0.072	0.082	0.161	0.131	0.360

Table A.1
Asset tickers.

U.S. futures		*Cryptocurrency*	
Ticker	Description	Ticker	Description
YM	Dow Mini	BTC	Bitcoin
NQ	E-Mini Nasdaq-100	ETH	Ethereum
ES	E-Mini S&P 500	XRP	Ripple
PA	Palladium	SOL	Solana
EW	E-Mini S&P 500 Midcap	BNB	Binance Coin
UB	U.S. Treasury Bond	ADA	Cardano
CL	Crude Oil WTI	DOGE	Doge Coin
VX	VIX		
RTY	E-Mini Russel 2000		

Table A.2
Number of stop-loss signals (positive) and average number of correct signals by ML models (true positive).

U.S. futures			*Cryptocurrency*		
Asset	Positive	Avg. true positive	Asset	Positive	Avg. true positive
YM	736	579.16	BTC	1245	814.85
NQ	728	498.97	ETH	1217	717.18
ES	728	511.71	XRP	1210	666.11
PA	662	343.05	SOL	1079	538.85
EW	661	474.60	BNB	1256	682.64
UB	681	415.75	ADA	1157	555.59
CL	693	317.32	DOGE	1156	627.82
VX	522	169.08			
RTY	718	487.16			

3.3.3. Classification results

We have demonstrated that the stop-loss adjusted labels are beneficial for implementing stop-loss strategies using various machine learning models. However, it is important to assess the impact of the proposed labeling scheme on the accuracy of prediction models. Therefore, we evaluate the performance of classification models with labels y_t and $\check{y}_t^{\delta}$, focusing on two aspects: (1) label difference, which measures the number of labels changed when $\check{y}_t^{\delta}$ is used instead of y_t, and (2) AUC[2] difference, which quantifies the difference of AUC between models trained with $\check{y}_t^{\delta}$ and y_t. More specifically, these can be expressed as follows:

Label difference: $\sum_t \mathbf{1}_{\check{y}_t^{\delta}=1} - \sum_t \mathbf{1}_{y_t=1}$

AUC difference: $AUC(\check{h}(X_t\ ), \check{y}_t^{\delta}\ \ ) - AUC(h(X_t\ ), y_t\ )$, where h is a predictor of ordinary labels, and $\check{\check{h}}$ is a predictor of stop-loss adjusted labels.

In Fig. 4, the upper row shows the label differences for different values of δ, and the lower row shows the AUC differences by aggregating all the results from the five prediction models (MLP, XG, RF, CB, and KNN). First, we can see that label differences become smaller as δ increases. This is natural because a larger δ means there is a smaller chance of hitting the stop-loss price. Second, all AUC differences are very close to 0, indicating that the prediction power is not much affected by stop-loss adjusted labels. Hence, these findings suggest that we can safely employ the stop-loss adjusted labels when implementing stop-loss strategies with machine learning models.

3.3.4. How does stop-loss adjusted labeling work inside machine learning model?

Now, we investigate the effect of the stop-loss adjusted labeling ($\check{y}_t^{\delta}$) within machine learning models. For this purpose, we use Shapely values (Lundeberg and Lee, 2017), which is one of the most popular tools for understanding the contributions of individual features in machine learning models. Table 3 presents a comparative analysis of the importance of input variables, measured in Shapely values, in classifying stop-loss adjusted labels $\check{y}_t^{\delta}$ for U.S. futures and cryptocurrencies. Each value in the table represents the average of the Shapely values calculated for all assets in either U.S. futures or cryptocurrencies. For each stop-loss threshold value (δ), the two most important variables were highlighted in bold. Additionally, some columns that show clear increasing or decreasing trends are emphasized with a gray gradation. While the table shows the results obtained from XGBoost, the other four models also exhibit similar tendencies.

For U.S. futures, z_{d15} and RSI are the two most important variables in most cases. However, as we tighten the stop-loss threshold, the importance of z_{high}, z_{low}, Volume MACD($MACD_V$), and OBV increases. On the other hand, the importance of z_{close}, z_{d5}, z_{d10}, and

[2] AUC refers to 'area under ROC (receiver operating characteristic) curve'. It is a measure of classification performance, known to be more appropriate than accuracy when labels are imbalanced (Ling et al., 2003; Shen, 2005).

[3] See Lee et al. (2023) for an overview of ML for asset management.

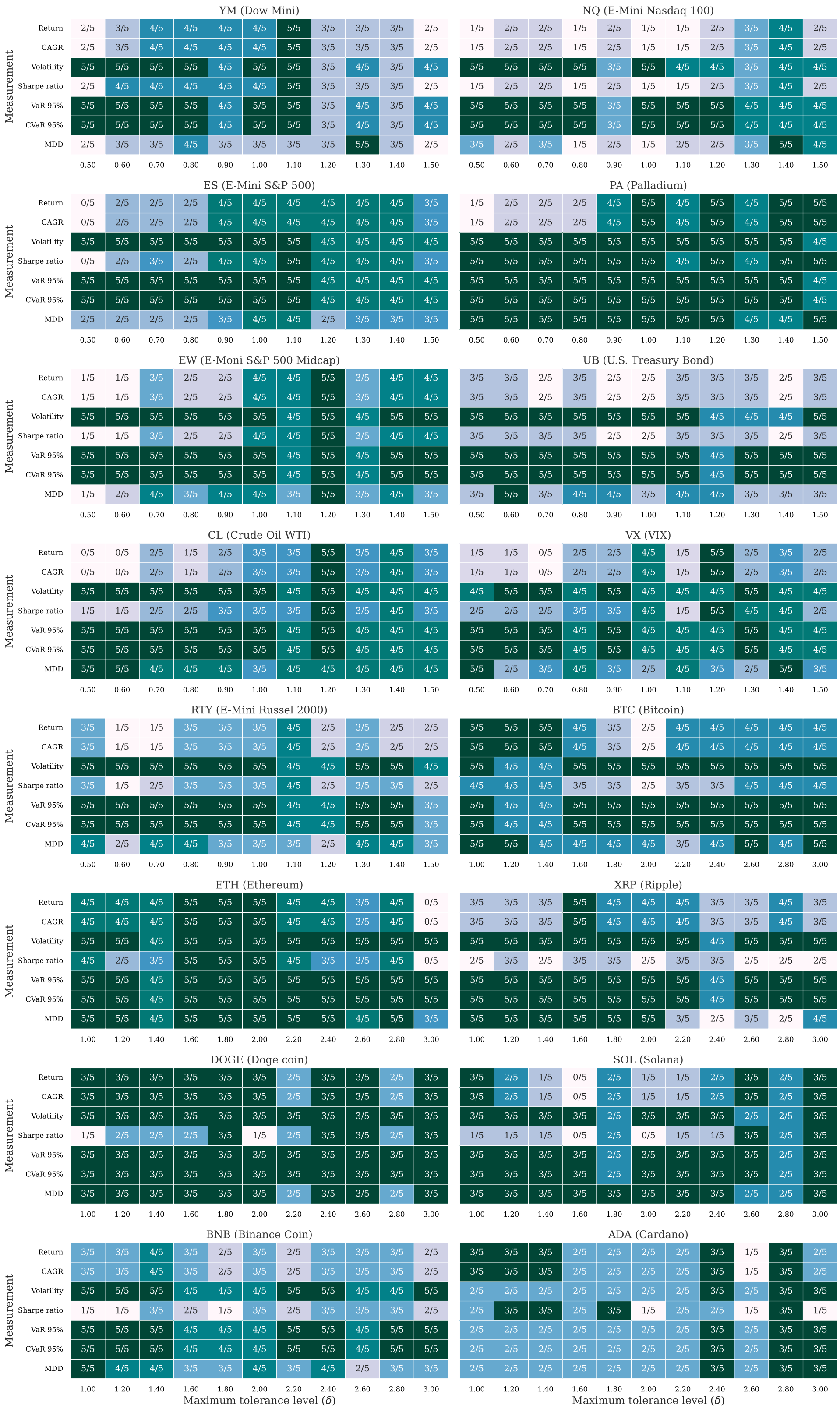

Fig. B.1. Win-matrix results for ST and MT on various U.S. futures and cryptocurrencies.

Table B.1
Detailed performance results of $\mathbf{ST}_\delta$ and $\mathbf{MT}_\delta$ with SVM.

Panel A. U.S. futures															
Asset	δ	Return		CAGR		Volatility		Sharpe ratio		VaR 95%		CVaR 95%		MDD	
		ST	MT	ST	MT	ST	MT	ST	MT	ST	MT	ST	MT	ST	MT
YM	1.20	**0.0889**	0.0313	**0.0750**	0.0265	**0.0600**	0.0676	**0.2807**	0.1143	**0.6200**	0.7000	**0.6700**	0.7000	**0.0757**	0.0974
	1.30	**0.1335**	0.0312	**0.1123**	0.0264	**0.0592**	0.0676	**0.4038**	0.1140	**0.6700**	0.7000	**0.6000**	0.7000	**0.0778**	0.0973
	1.40	**0.2609**	0.0312	0.2174	**0.0264**	**0.0661**	0.0676	**0.6527**	0.1140	**0.6000**	0.7000	**0.6200**	0.7000	**0.0914**	0.0974
NQ	1.20	**0.2809**	-0.0020	**0.2339**	-0.0017	0.1024	**0.1014**	**0.4778**	0.0472	**1.0400**	1.0500	**0.9800**	1.0500	**0.1561**	0.2523
	1.30	**0.2123**	-0.0092	**0.1776**	-0.0078	0.1053	**0.1019**	**0.3754**	0.0349	**0.9800**	1.0500	1.0800	**1.0500**	**0.1582**	0.2491
	1.40	**0.0915**	-0.0062	**0.0772**	-0.0053	**0.0949**	0.1017	**0.2104**	0.0400	**1.0800**	1.0500	**1.0400**	1.0500	**0.2227**	0.2545
ES	1.20	**0.0706**	-0.0361	**0.0596**	-0.0308	**0.0736**	0.0775	**0.2007**	-0.0451	**0.7600**	0.8000	**0.7600**	0.8000	**0.1115**	0.1625
	1.30	**0.1812**	-0.0328	**0.1518**	-0.0279	**0.0696**	0.0773	**0.4575**	-0.0375	**0.7600**	0.8000	**0.7100**	0.8000	**0.1735**	0.1671
	1.40	**0.1009**	-0.0308	**0.0850**	-0.0262	**0.0742**	0.0771	**0.2658**	-0.0330	**0.7100**	0.8000	**0.7600**	0.8000	**0.0961**	0.1642
PA	1.20	**1.1759**	0.5450	**0.9347**	0.4466	**0.1572**	0.1856	**0.9466**	0.5024	**1.5700**	1.8900	**1.6300**	1.8900	**0.2054**	0.3177
	1.30	**1.0307**	0.6027	**0.8245**	0.4924	**0.1646**	0.1864	**0.8375**	0.5356	**1.6300**	1.8900	**1.6500**	1.8900	**0.1971**	0.3202
	1.40	**1.1641**	0.5570	**0.9257**	0.4562	**0.1633**	0.1864	**0.9115**	0.5084	**1.6050**	1.8900	**1.5700**	1.8900	**0.2264**	0.3303
EW	1.20	**0.1541**	0.1210	**0.1293**	0.1018	**0.0886**	0.0922	**0.3606**	0.2883	**0.9000**	0.9400	**0.9000**	0.9500	**0.1490**	0.1687
	1.30	0.0548	**0.1117**	0.0463	**0.0940**	**0.0869**	0.0928	0.1636	**0.2696**	**0.9000**	0.9500	**0.8900**	0.9500	**0.1259**	0.1652
	1.40	0.1104	**0.1163**	0.0930	**0.0979**	**0.0878**	0.0925	0.2772	**0.2790**	**0.8900**	0.9500	**0.9000**	0.9400	0.1914	0.1721
UB	1.20	**0.3461**	0.3324	**0.2870**	0.2759	0.1335	0.1335	**0.4512**	0.4377	1.3600	1.3600	**1.3500**	1.3600	0.1180	**0.1056**
	1.30	0.2493	**0.3324**	0.2080	**0.2759**	0.1333	0.1335	0.3527	**0.4377**	**1.3500**	1.3600	1.3600	1.3600	0.1460	**0.1056**
	1.40	0.3105	**0.3324**	0.2580	**0.2759**	**0.1327**	0.1335	0.4173	**0.4377**	1.3600	1.3600	1.3600	1.3600	0.1576	**0.1056**
CL	1.20	**0.6151**	0.6104	**0.5022**	0.4985	**0.1456**	0.1514	**0.6427**	0.6201	**1.4700**	1.5300	1.6000	**1.5600**	0.2664	**0.2400**
	1.30	**0.6480**	0.6373	**0.5480**	0.5197	0.1647	**0.1532**	**0.6479**	0.6337	1.6000	1.5600	1.7000	**1.5500**	0.2224	**0.2211**
	1.40	0.3062	**0.5322**	0.2545	**0.4365**	0.1568	**0.1536**	0.3724	0.5576	1.7000	1.5500	**1.4700**	1.5300	0.3505	**0.2111**
VX	1.20	**0.1736**	0.0259	**0.1455**	0.0220	**0.1443**	0.1516	**0.2668**	0.1046	**1.4800**	1.5600	**1.4100**	1.5600	0.2492	0.2373
	1.30	-0.0034	**0.0113**	-0.0029	**0.0096**	**0.1004**	0.1532	0.0439	**0.0887**	**1.4100**	1.5600	**1.0400**	1.5800	**0.1822**	0.2361
	1.40	-0.0074	**0.0471**	-0.0063	**0.0399**	**0.1363**	0.1518	0.0577	**0.1285**	**1.0400**	1.5800	**1.4800**	1.5600	**0.1986**	0.2432
RTY	1.20	**0.2097**	0.0106	**0.1754**	0.0090	0.0937	0.0937	**0.4058**	0.0667	**0.9600**	0.9700	**0.8800**	0.9700	0.1591	**0.1517**
	1.30	**0.1187**	0.0090	**0.0999**	0.0076	**0.0907**	0.0938	**0.2638**	0.0637	**0.8800**	0.9700	**0.9300**	0.9700	0.1848	**0.1566**
	1.40	**0.0817**	0.0101	**0.0690**	0.0086	**0.0859**	0.0937	**0.2045**	0.0659	**0.9300**	0.9700	**0.9600**	0.9700	**0.1482**	0.1517

Panel B. Cryptocurrency															
Asset	δ	Return		CAGR		Volatility		Sharpe ratio		VaR 95%		CVaR 95%		MDD	
		ST	MT	ST	MT	ST	MT	ST	MT	ST	MT	ST	MT	ST	MT
BTC	2.40	**-0.0995**	-0.2833	**-0.0846**	-0.2448	**0.1486**	0.1764	**0.0058**	-0.0949	**1.5400**	1.8300	**1.5400**	1.8400	**0.4549**	0.6033
	2.60	**0.1024**	-0.2678	**0.0856**	-0.2310	**0.1567**	0.1772	**0.1384**	-0.0819	**1.5400**	1.8400	**1.6200**	1.8400	**0.5436**	0.6037
	2.80	**0.2133**	-0.2799	**0.1770**	-0.2418	**0.1494**	0.1768	**0.1998**	-0.0917	**1.6200**	1.8400	**1.5400**	1.8300	**0.4697**	0.5924
ETH	2.40	**-0.1674**	-0.2597	**-0.1431**	-0.2239	**0.1741**	0.2403	-0.0152	**-0.0018**	**1.8000**	2.4900	**1.8600**	2.5000	0.7217	**0.6727**
	2.60	**0.1949**	-0.0102	**0.1619**	-0.0086	**0.1913**	0.2443	**0.1855**	0.1174	**1.8600**	2.5000	**1.9700**	2.5200	**0.6163**	0.7071
	2.80	-0.2250	**-0.1009**	-0.1934	**-0.0858**	**0.1787**	0.2415	-0.0491	**0.0774**	**1.9700**	2.5200	**1.8000**	2.4900	**0.6304**	0.6575
XRP	2.40	-0.2351	**0.1450**	-0.2022	**0.1209**	**0.1719**	0.2040	-0.0652	**0.1652**	**1.7900**	2.1000	**1.3300**	2.1000	**0.4376**	0.4914
	2.60	-0.2204	**0.1765**	-0.1894	**0.1469**	**0.1336**	0.2057	-0.1138	**0.1785**	**1.3300**	2.1000	**1.3900**	2.1200	0.5722	**0.5182**
	2.80	-0.2823	**0.2103**	-0.2440	**0.1746**	**0.1270**	0.2044	-0.1895	**0.1916**	**1.3900**	2.1200	**1.7900**	2.1000	**0.4261**	0.5131
SOL	2.40	**0.1514**	-0.3152	**0.1262**	-0.2732	**0.2941**	0.2994	**0.1913**	0.0248	**3.0200**	3.1000	**2.8900**	3.1300	**0.7616**	0.8442
	2.60	**0.5448**	-0.4153	**0.4428**	-0.3639	**0.3022**	0.3043	**0.2880**	-0.0208	**2.8900**	3.1300	**3.1000**	3.1600	**0.6502**	0.8286
	2.80	**-0.2528**	-0.3692	**-0.2178**	-0.3219	**0.2797**	0.3019	**0.0372**	0.0008	**3.1000**	3.1600	**3.0200**	3.1000	**0.6432**	0.8503
BNB	2.40	**-0.3514**	-0.7020	**-0.3058**	-0.6396	**0.1653**	0.2443	**-0.1717**	-0.3591	**1.7200**	2.5700	**1.8700**	2.5900	**0.5811**	0.7068
	2.60	**-0.3370**	-0.6259	**-0.2928**	-0.5634	**0.1854**	0.2507	**-0.1226**	-0.2556	**1.8700**	2.5900	**1.9300**	2.6200	**0.5719**	0.7330
	2.80	**-0.2618**	-0.6383	**-0.2258**	-0.5757	**0.1800**	0.2475	**-0.0741**	-0.2755	**1.9300**	2.6200	**1.7200**	2.5700	**0.6099**	0.7376
ADA	2.40	**-0.2749**	-0.6091	**-0.2374**	-0.5470	**0.1664**	0.2460	**-0.1052**	-0.2521	**1.7300**	2.5700	**1.9400**	2.6000	**0.3602**	0.6896
	2.60	**-0.4954**	-0.5808	**-0.4382**	-0.5195	**0.2324**	0.2495	**-0.1741**	-0.2179	**1.9400**	2.6000	**2.4200**	2.6100	**0.3553**	0.6727
	2.80	**-0.2581**	-0.6310	**-0.2225**	-0.5684	**0.1866**	0.2481	**-0.0635**	-0.2703	**2.4200**	2.6100	**1.7300**	2.5700	**0.5345**	0.6539
DOGE	2.40	**-0.0995**	-0.2833	**-0.0846**	-0.2448	**0.1486**	0.1764	**0.0058**	-0.0949	**1.5400**	1.8300	**1.5400**	1.8400	**0.4549**	0.6033
	2.60	**0.1024**	-0.2678	**0.0856**	-0.2310	**0.1567**	0.1772	**0.1384**	-0.0819	**1.5400**	1.8400	**1.6200**	1.8400	**0.5436**	0.6037
	2.80	**0.2133**	-0.2799	**0.1770**	-0.2418	**0.1494**	0.1768	**0.1998**	-0.0917	**1.6200**	1.8400	**1.5400**	1.8300	**0.4697**	0.5924

Price MACD($MACD_P$) decreases. Hence, we can see that incorporating stop-loss adjusted labels causes machine learning models to focus more on high and low prices, as well as volumes. As for cryptocurrencies, the trends for z_{high} and z_{low} are not as clear as U.S. futures, but we can still see an increase in the importance of Volume MACD ($MACD_V$) and OBV as we make the stop-loss threshold tighter.

4. Conclusion

While many researchers and practitioners are trying to use machine learning models to predict prices of financial assets, most of them are not perfectly aligned with actual implementations. To make prediction and decision-making in trading problems more aligned, we propose integrating stop-loss strategies into prediction models through a stop-loss adjusted labeling scheme. Numerical experiments with five different machine learning models suggest that simply adjusting labels to incorporate stop- can significantly reduce risk. That is, our study can help traders to achieve better implementations of stop-loss strategies with machine learning models.

CRediT authorship contribution statement

Yoontae Hwang: Writing – original draft, Writing – review & editing, Data curation, Formal analysis, Visualization, Software. **Junpyo Park:** Data curation. **Yongjae Lee:** Writing – original draft, Writing – review & editing, Methodology, Visualization, Funding acquisition. **Dong-Young Lim:** Writing – original draft, Writing – review & editing, Methodology, Conceptualization, Funding acquisition.

Data availability

Data is availabel at: https://github.com/Yoontae6719/Stop-loss-adjusted-labels.

Acknowledgement

Yongjae Lee acknowledge that this research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2022R1I1A4069163). Dong-Young Lim acknowledge that this work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2020-0-01336, Artificial Intelligence Graduate School Program (UNIST)) and by the 2023 Research Fund (1.230035.01) of UNIST (Ulsan National Institute of Science & Technology).

Appendix

Appendix A. Assets

Table A.1
Table A.2

Appendix B. Additional Experiment Results

Fig. B.1
Table B.1

References

Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M., 2019. Optuna: a next-generation hyperparameter optimization framework. *ACM SIGKDD international conference on knowledge discovery & data mining*. 2623–2631.
Dai, B., Marshall, B.R., Nguyen, N.H., Visaltanachoti, N., 2021. Risk reduction using trailing stop-loss rules. Int. Rev. Finance 21 (4), 1334–1352.
Feng, F., Chen, H., He, X., Ding, J., Sun, M., Chua, T.S., 2019. Enhancing stock movement prediction with adversarial training. IJCAI 5843–5849.
Gao, T., Chai, Y., 2018. Improving stock closing price prediction using recurrent neural network and technical indicators. Neural Comput. 30 (10), 2833–2854.
Gonçalves, R., Ribeiro, V.M., Pereira, F.L., Rocha, A.P., 2019. Deep learning in exchange markets. Inform. Econ. Policy 47, 38–51.
Ghosh, P., Neufeld, A., Sahoo, J.K., 2022. Forecasting directional movements of stock prices for intraday trading using LSTM and random forests. Finance Res. Lett. 46, 102280.
Gu, S., Kelly, B., Xiu, D., 2020. Empirical asset pricing via machine learning. Rev. Financ. Stud. 33 (5), 2223–2273.
Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. J. Finance 45 (3), 881–898.
Jegadeesh, N., 1991. Seasonality in stock price mean reversion: evidence from the US and the UK. J. Finance 46 (4), 1427–1444.
Ju, G., Kim, K.K., Lim, D.Y., 2019. Learning multi-market microstructure from order book data. Quant. Finance 19 (9), 1517–1529.
Kaminski, K.M., Lo, A.W., 2014. When do stop-loss rules stop losses? J. Financ. Markets 18, 234–254.
Lee, Y., Thompson, R.J., Kim, J.H., Kim, W.C., Fabozzi, F.A., 2023. An overview of machine learning for asset management. J. Portf. Manag. https://doi.org/10.3905/jpm.2023.1.526.
Lei, A.Y., Li, H., 2009. The value of stop loss strategies. Financ. Serv. Rev. 18 (1), 23–51.

Han, Y., Zhou, G., Zhu, Y., 2016. Taming momentum crashes: A simple stop-loss strategy. SSRN Electronic J. https://doi.org/10.2139/ssrn.2407199.
Ling, C.X., Huang, J., Zhang, H., 2003. AUC: a statistically consistent and more discriminating measure than accuracy, International Joint Conference on Artificial Intelligence (IJCAI). 519–524.
Liu, J., Serletis, A., 2019. Volatility in the cryptocurrency market. Open Econ. Rev. 30, 779–811.
Lo, A.W., Remorov, A., 2017. Stop-loss strategies with serial correlation, regime switching, and transaction costs. J. Financ. Markets 34, 1–15.
Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst 30.
Nelson, D.M.Q., Pereira, A.C.M., De Oliveira, R.A., 2017. Stock market's price movement prediction with LSTM neural networks. *International Joint Conference on Neural Networks (IJCNN)*. 1419–1426.
Patel, J., Shah, S., Thakkar, P., Kotecha, K., 2015. Predicting stock market index using fusion of machine learning techniques. Expert Syst. Appl. 42 (4), 2162–2172.
Rundo, F., Trenta, F., di Stallo, A.L., Battiato, S., 2019. Grid trading system robot (gtsbot): a novel mathematical algorithm for trading fx market. Appl. Sci. 9 (9), 1796.
Santos, A.A.P., Torrent, H.S., 2022. Markowitz meets technical analysis: building optimal portfolios by exploiting information in trend-following signals. Finance Res. Lett. 49, 103063.
Shen, Y., 2005. Loss Functions For Binary Classification and Class Probability Estimation. University of Pennsylvania. PhD thesis.
Sirignano, J.A., 2019. Deep learning for limit order books. Quant. Finance 19 (4), 549–570.
Sirignano, J., Cont, R., 2019. Universal features of price formation in financial markets: perspectives from deep learning. Quant. Finance 19 (9), 1449–1459.
Sun, X., Liu, M., Sima, Z., 2020. A novel cryptocurrency price trend forecasting model based on LightGBM. Finance Res. Lett. 32, 101084.
Yoo, J., Soun, Y., Park, Y.C., & Kang, U. 2021. Accurate multivariate stock movement prediction via data-axis transformer with multi-level contexts. *In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining,* 2037–2045.
Yeh, W.C., Hsieh, Y.H., & Huang, C.L. (2022). Newly developed flexible grid trading model combined ANN and SSO algorithm. arXiv preprint arXiv:2211.12839.