

메타모형을 이용한 옵션 양매도 전략의 마켓타이밍 예측연구

김찬수
한국펀드평가

A Market Timing Prediction of Option Short Straddle Strategy Using a Meta Model

Chansu Kim
Korea Fund Ratings

요약 본 연구는 EBM(Explainable Boosting Machine)을 이용하여 옵션 양매도 전략의 장 중 포지션 진입 여부를 결정하는 메타모형을 제안하고 투자 성과를 분석하였다. 실험에는 2004년부터 2018년까지의 코스피 200 주가지수옵션 자료와 시장 외생 변수를 활용했고, 95개의 특성을 가공한 후, 특성 중요도, T-test, 수동선택을 통해 22개의 특성을 선택하여 모형을 학습하였다. 실험 결과는 다음과 같다. 첫째, 메타모형의 성능은 정밀도 0.6404, 재현율 0.7828, F1 점수 0.7045로, 투자 전략의 성공 여부를 약 70%의 확률로 예측함을 보여주었다. 둘째, 예측에 영향을 미치는 주요 특성변수로는 요일, VIX 10일 표준편차, 항생지수 3일 표준편차로 확인되었다. 셋째, 제안된 모형으로 필터링한 전략의 샤프 비율은 벤치마크를 상회하는 결과를 보였다. 이 결과는 메타모형을 통한 거래 필터링이 효과적임을 보여준다.

Abstract This study proposed a meta-model for determining the entry positions for an option strangle strategy using the Explainable Boosting Machine (EBM) and its investment performance. Using the KOSPI 200 Index Options data and exogenous market variables from 2004 to 2018, 95 features were engineered, and 22 were selected through feature importance, T-tests, and manual selection to train the model. The key findings were as follows. First, the performance of the meta-model, with a precision of 0.6404, a recall of 0.7828, and an F1 score of 0.7045, demonstrates the capability of the model to predict the success of the investment strategy with approximately 70% probability. Second, the key feature variables influencing the predictions were confirmed as the day of the week, the 10-day standard deviation of VIX, and the three-day standard deviation of the Hang Seng Index Third, the Sharpe ratio of the strategy filtered using the proposed model exceeded the benchmark. These results highlight the effectiveness of transaction filtering through the meta-model.

Keywords : Machine Learning, Option Trading, Short Straddle, Market Timing, Meta Labeling

1. 서론

옵션(option)은 특정 기간 특정 가격으로 기본 자산을 매입하거나 매도할 수 있는 권리를 제공하는 파생상품이다. 권리는 매수자가 매도자에게 프리미엄을 지불함으로써 얻게 되며, 이 프리미엄이 옵션 계약의 핵심 가치

를 결정한다.

옵션은 그 독특한 구조 덕분에 다양한 방식으로 활용되고 있다. 한 가지 주요 활용 방법은 헷징(hedging)이다. 투자자들은 주식 같은 위험자산의 가격 변동으로부터 포트폴리오를 보호하는 목적으로 옵션을 사용할 수 있다 [1]. 또 다른 활용 방법은 옵션 매매(option trading)다.

*Corresponding Author : Chansu Kim(Korea Fund Ratings)

email: cskim@kfr.co.kr

Received July 25, 2023

Accepted October 6, 2023

Revised August 16, 2023

Published October 31, 2023

특히 옵션은 방향성(directionality)과 변동성(volatility) 예측에 기반을 둔 다양한 투자 전략을 가능하게 하므로, 투자자들에게 다양한 선택의 여지를 제공한다.

대표적인 옵션 매매 전략에는 양매도라고 부르는 'short straddle'이 있다. 양매도는 기초자산의 가격이 크게 움직이지 않고 시간이 지나거나 변동성이 감소하면 이익을 얻을 수 있어, 방향성 매매보다 승률이 높은 것이 특징이다. 그러나 이 전략은 기초자산의 가격이 한쪽으로 크게 오르거나 변동성이 증가하는 경우에는 큰 손실을 볼 수 있다.

옵션투자자들은 기술적 분석, 펀더멘탈 분석 등을 기반으로 전략을 만들어 손실을 최소화하기 위해 노력했다. 하지만 이런 방법들은 높은 차원의 데이터를 처리하거나 복잡한 비선형 관계를 탐색하는 데는 한계가 있다. 이에 최근에는 옵션 시장에서 최적의 전략을 찾는 데 기계학습이 잠재력을 보여주고 있다[2-5].

본 연구는 기계학습을 활용해 양매도 전략의 진입 여부를 결정하는 메타모형(meta model)을 제안하고, 그 수익성을 분석한다. 메타모형의 학습을 위해 2004년 1월 5일부터 2018년 12월 28일까지의 VIX, 환율, 미국 주가지수 등의 외생 변수와 코스피200 주가지수 옵션의 증가, 시가, 미결제약정 등 옵션 데이터를 활용했다.

본 연구의 주요한 차별점은 95개의 외생 변수를 특성 중요도(feature importance), T-test, 수동(manual)선택을 통해 분석한 후 22개의 주요 특성을 선택한 것과 양매도 전략 결과의 크기를 예측하는 메타모형을 구축해 성과를 개선한 점이다. 이는 베팅 방향(betting side)을 예측한 기존 양매도 연구와 비교했을 때 새로운 접근법을 제시한다.

2. 이론적 배경

2.1 옵션 양매도 전략

양매도는 동일한 만기일과 행사 가격을 가진 콜 옵션과 풋 옵션을 동시에 매도하여 수익을 얻는 전략이다. 양매도는 기초자산의 가격이 크게 움직이지 않으면 옵션의 시간 가치 소멸에 따라 수익을 얻을 수 있다. Fig. 1은 옵션 양매도의 수익구조 매커니즘을 보여주고 있다. 이 전략은 변동성이 낮고 안정적인 상황에서 가장 효과적이다. 그러나 만약 기초자산의 가격이 크게 변동하면, 감마 리스크(gamma risk)에 의해 매도한 옵션 중 하나가 크게 가치를 가지게 되어 큰 손실을 볼 수 있다.

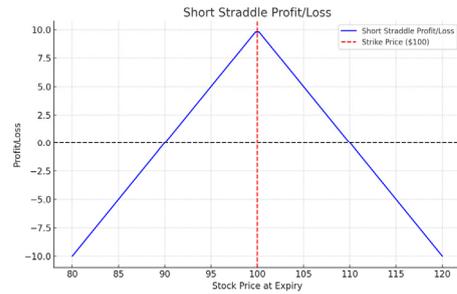


Fig. 1. Option short straddle strategy

2.2 일별 양매도 전략과 변동성

일반적인 옵션 매도전략은 옵션의 시간 가치 감소를 통해 수익을 추구한다. 만기일이 다가올수록 옵션 행사 가능성이 줄어들어 가격 하락에 대한 프리미엄을 취할 수 있기 때문이다. 반면, 일별 양매도 전략은 하루 동안의 변동성 축소를 통해 이익을 얻는 것을 목표로 한다. 즉, 하루 중 기초자산의 변동성이 하락하면, 그 결과로 매도한 옵션의 가치가 감소하면서 수익을 얻을 수 있다.

변동성은 하루 동안 하락하는 특성이 있다[6]. 이는 거래자들이 전일 장 마감 이후 발생한 정보를 장 시작과 동시에 처리하는 경향 때문이다[7]. 시장이 시작될 때는 정보에 대한 거래자들의 반응으로 높은 변동성을 보이지만, 시간이 흐르면 해당 정보가 가격에 반영되면서 변동성은 감소한다. 이처럼 하루 중 변동성이 감소하는 경향 덕분에 일별 양매도 전략은 높은 샵피비율을 보인다[8].

2.3 메타 라벨링(meta labeling)

De Prado(2018)에 의해 도입된 메타 라벨링은 1차 모형(primary model)의 예측 결과를 외생변수들과 함께 학습하여 2차 모형(secondary model)을 생성하고, 이를 통해 베팅의 크기(betting size)를 결정하는 방법이다[9].

Fig. 2는 메타 라벨링 프로세스를 보여준다. 1차 모형은 매수와 매도와 같은 트레이딩 신호를 예측하지만, 이 신호들은 항상 정확하지 않다. 이 문제를 해결하기 위해 2차 모형인 메타 모형이 사용된다. 메타 모형은 1차 모형에서 생성된 신호의 신뢰도를 평가하여 해당 신호를 따를지 결정한다. 이 과정을 통해 메타 모형은 1차 모형의 정밀도를 보정하고 전반적인 트레이딩 성능을 향상시킨다. 이처럼 메타 라벨링은 베팅의 방향과 크기의 예측을 분리함으로써, 과적합(overfitting)을 예방하는 동시에 기계학습 모형을 더욱 이해하기 쉽게 만들어 준다는 장

점이 있다[10].

작은 베팅에서는 높은 정확도를, 큰 베팅에서는 낮은 정확도를 보일 경우, 그 거래는 성공하기 어렵다. 따라서 거래의 기회를 발견하는 것뿐만 아니라, 적절한 베팅 크기를 설정하는 것이 매우 중요하다. 이를 위해 올바른 베팅 크기를 결정하는 메타모형을 개발하는 것은 큰 의미가 있다.

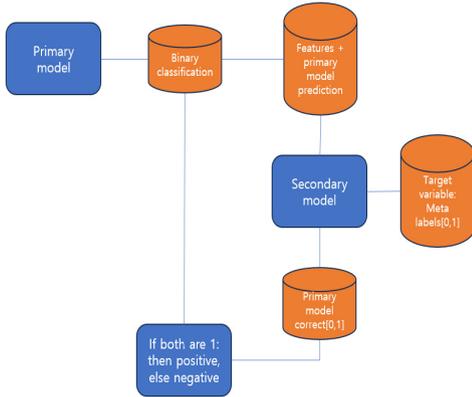


Fig. 2. Architecture of meta labeling

2.4 Explainable Boosting Machine(EBM)

EBM은 해석 가능성을 최우선으로 하는 머신러닝 모형이다[11]. 이 모형은 일반화 가법 모형(Generalized Additive Model, GAM)의 한 형태인데, 다음과 같은 특징을 갖고 있다.

첫째, EBM은 배깅(bagging), 그래디언트 부스팅(gradient boosting)과 같은 현대적인 머신러닝 기법으로 특정 함수를 학습하여 우수한 성능을 보인다. 둘째, EBM은 한 변수의 효과가 다른 변수의 수준에 따라 달라지는 상호작용(interaction)을 자동 감지함으로써 모형의 해석 가능성을 높인다.

Table 1은 EBM, 랜덤 포레스트(RF), 부스팅 모형의 주요 특징을 비교하고 있다. EBM은 각 특성의 개별적인 효과플롯(effect plot)을 제공하기 때문에 모형의 예측을 명확하게 해석할 수 있으며, 높은 성능과 빠른 학습 속도를 동시에 보여준다.

Table 1. Comparison of EBM, RF, Boosting

Feature	EBM	RF	Boosting
Interpretability	High	Medium	Low
Performance	High	High	High
Interaction Detection	Yes	Limited	Limited
Training Speed	Fast	Medium	Slow

Fig. 3은 성인 소득 데이터를 EBM에 적용한 예시이다. 위쪽 그래프에서는 성인 소득에 가장 큰 영향을 주는 변수가 나이(age)임을 알 수 있고, 아래쪽 그래프에서는 20대부터 50대까지는 소득이 증가하는 경향이 있으나, 60대 이후로는 소득이 감소하는 것을 확인할 수 있다.



Fig. 3. Example of EBM function graph

3. 데이터와 제안모형 소개

3.1 데이터 소개

Table 2는 사용된 원본 데이터(raw data)를 보여주고 있다. 이 데이터는 2004년 6월 28일부터 2018년 12월 28일까지의 기간 동안 수집된 것이며, 이를 가공하여 학습에 필요한 데이터를 생성했다. Table 3은 가공된 학습 데이터를 나타내는데, 이 데이터는 각 지수의 일별 수익률, 3일, 5일, 10일의 이격도 및 표준편차, 그리고 옵션 데이터를 포함하여 총 95개의 데이터로 구성되어 있다. 이격도(disparity ratio)는 이동평균선과의 이격도를 나타내는 지표로 (1)과 같이 산출하였다.

$$Disparityratio = \left(\frac{Close}{Moving\ Average} - 1 \right) \quad (1)$$

Table 2. Explanation of raw data

Item	Description	Source
Copen	Call Open	Korea Exchange
Popen	Put Open	Korea Exchange
PrevCclose	Previous Call Close	Korea Exchange
PrevPclose	Previous Put Close	Korea Exchange
PrevCOI	Previous Call Open interest	Korea Exchange
PrevPOI	Previous Put Open interest	Korea Exchange
KPopen	KOSPI Open	Daishin Securities
KPprevclose	KOSPI Previous Close	Daishin Securities
VIX	CBOE Volatility Index	Investing.com
SHANG	Shanghai	Daishin Securities
SPX	S&P 500	Daishin Securities
COMP	NASDAQ	Daishin Securities
HSCE	Hang Seng	Daishin Securities
USD/KRW	Exchange Rate	Investing.com
KQ11	KOSDAQ	Daishin Securities
KS11	KOSPI	Daishin Securities
99915	Gold	Daishin Securities
99925	US Bond 10Y	Daishin Securities
99944	US Bond 2Y	Daishin Securities
CM@DUB1M	Crude Oil	Daishin Securities

Table 3. Processed data for training

Item	Description
Open ratio	Copen/Popen
Prev OI Diff	PrevPOI-PreVCOI
Day of month	Date
Weekday	Mon:0, Tues:1 Wed:2, Thurs:3, Fri:4
Gap	KP200 open/prevclose-1
VIX	Daily return, 3, 5 10 days of disparity ratio and standard deviation
SHANG	
SPX	
COMP	
HSCE	
USD/KRW	
KQ11	
KS11	
99915	
99925	
99944	
CM@DUB1M	

3.2 일별 양매도 데이터

본 연구에서는 매일 9시 1분에 각 옵션 증가격(At the money, ATM)에 양매도를 진입하여 증가에 청산하는 전략을 1차 모형으로 결정하고, 학습에 이용했다. 일별

수익률은 (2)와 같은 방식으로 산출했다.

$$Dailyreturn = - \frac{(Cclose + Pclose)}{(Cclose_{901} + Pclose_{901})} - 1 \quad (2)$$

Fig. 4는 일별 양매도 전략의 수익률 히스토그램을 보여주고, Table 4는 표본데이터 3,712개에 대한 기초통계량을 보여주고 있다. 평균은 0.0061, 표준편차는 0.1691로, 데이터들이 평균값 주변에 넓게 분포하고 있다. 또한, 왜도는 -2.7065로 좌측으로 치우쳐있고, 첨도는 27.6171로 극단적인 값들이 자주 등장함을 나타낸다.

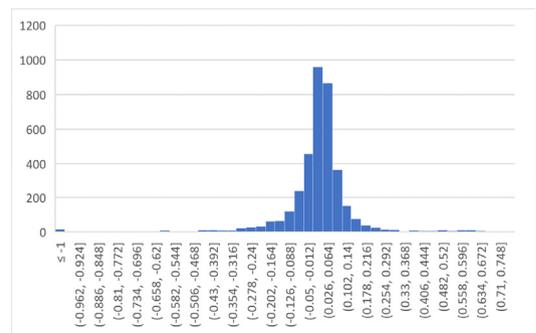


Fig. 4. Histogram of daily returns for short straddle strategy

Table 4. Summary statistics of daily short straddle returns

Statistics	Value
Average	0.0061
Standard deviation	0.1691
Skewness	-2.7065
Kurtosis	27.6171
Min	-1.9014
Max	0.7846
Count	3712

3.3 제안모형

메타모형은 3.1절에 제시된 변수를 특성변수로 사용한다. 타겟(target) 변수는 진입 후 수익이 발생한 경우 1, 아닌 경우 0으로 설정한다. 모형의 결과는 0에서 1 사이의 실수로 표현된다. 이 값을 통해 매매 비중을 결정하거나, 매매 전략의 진입 여부를 판단할 수 있다.

전체적인 구성은 Fig. 5에 표현되었으며, 구체적인 학습 과정은 다음과 같다.

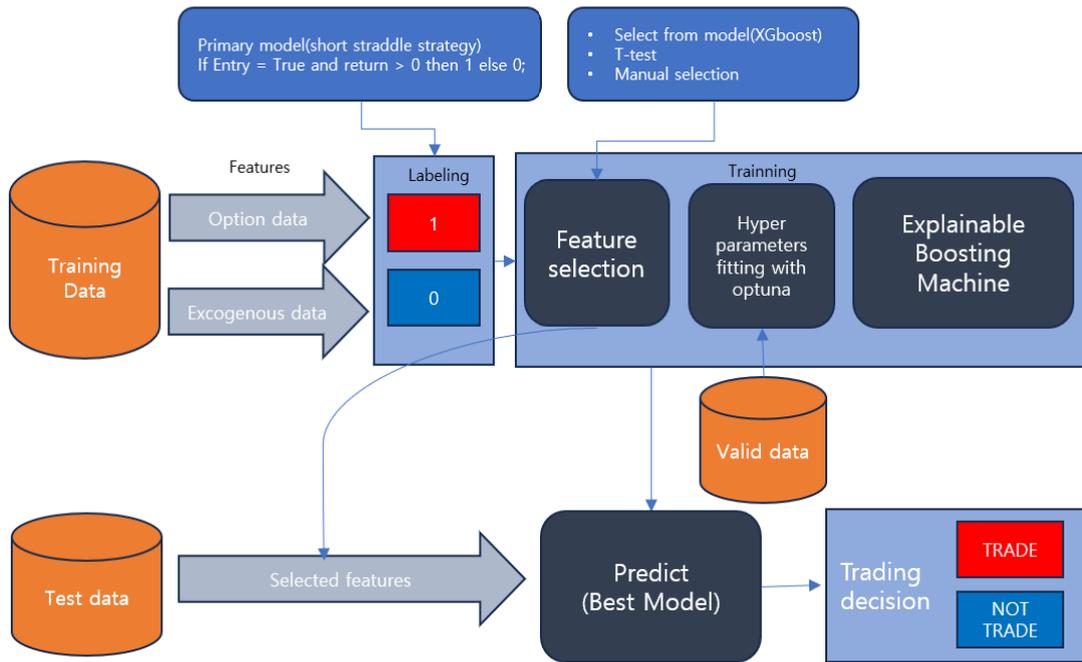


Fig. 5. Flowchart of suggested meta-model

먼저 2004년 6월 28일부터 2018년 12월 28일을 학습(64%), 검증(16%), 테스트(20%) 데이터로 나누었다. 이때 De Prado(2018)의 방법에 따라, 표본 구간의 경계값이 다른 표본에 미치는 영향을 최소화하기 위해, TSCV모듈의 'gap_train_test_split' 기능을 사용하여 각 구간 간에 공백을 설정했다[12].

학습 데이터의 라벨링 결과 1의 개수는 1986개 0의 개수는 1218개로 불균형이 나타났다. 이를 해결하기 위해 imblearn 모듈의 'RandomOverSampler'를 이용해 0의 개수를 증가시켰다[13].

그다음에는 95개의 특성변수 중에서, 결괏값에 영향을 미치는 변수들을 선택하는 작업을 진행했다. 이 과정에서 sklearn의 'SelectFromModel', 통계적 검정방법인 T-test를 사용했고, 일부 변수는 수동으로 선택했다. 이렇게 해서 최종적으로 22개의 특성변수를 선정했고, 이 변수들은 Table 5에 정리되어 있다.

모형의 하이퍼파라미터는 'Optuna'를 통해 선택했다 [14]. 목표 변수는 F1 점수로 설정했고, 100번의 시도를 통해 하이퍼파라미터를 결정했다. Table 6은 선택된 파라미터(best parameter)를 보여주고 있다.

Table 5. Selected features

Item	Description	Date
99944_Std_5	US 2Y Bond 5-day Std	T-1
COMP_Std_10	NASDAQ 10-day Std	T-1
Day of month	Date	T
Gap	KOSPI Gap	T
HSCE_Std_10	Hang Seng 10-day Std	T-1
HSCE_Std_3	Hang Seng 3-day Std	T-1
HSCE_Std_5	Hang Seng 5-day Std	T-1
KQ11_Disparity_10	KOSDAQ 10-day Disparity	T-1
KQ11_Disparity_3	KOSDAQ 3-day Disparity	T-1
KS11_Disparity_10	KOSPI 10-day Disparity	T-1
KS11_Disparity_5	KOSPI 5-day Disparity	T-1
Open ratio	Put, Call Open price Ratio	T
PrevClose	Previous Put Close	T-1
Prev OI diff	Previous Open Interest Diff.	T-1
PrevPOI	Previous Put Open Interest	T-1
SHANG_Std_10	Shanghai 10-day Std	T-1
SPX_Std_10	S&P 500 10-day Std	T-1
VIX_Disparity_3	Volatility Index 3-day Disparity	T-1
VIX_Disparity_5	Volatility Index 5-day Disparity	T-1
VIX_Std_10	Volatility Index 10-day Std	T-1
VIX_Std_5	Volatility Index 5-day Std	T-1
Weekday	Day of the week	T

Table 6. Best parameters of final model

Parameter	Value
Max_bins	142
Max_interaction_bins	19
Validation_size	0.18
Learning_rate	0.08
Min_samples_leaf	4
Max_leaves	2
Max_rounds	4987

4. 실험결과 분석

4.1 예측 결과 회귀분석

최종 모형 예측력의 통계적 유의성을 확인하기 위해 (3)과 같이 회귀분석을 진행했다. 여기서 독립 변수는 모형이 테스트셋(test set)을 예측한 0과 1 사이의 상승 확률을 사용하였고, 종속 변수는 일별 양매도 전략의 성과를 적용하였다.

$$Y_t = \beta_0 + \beta_1 X_t + \epsilon_t \quad (3)$$

Table 7은 회귀분석 결과를 보여준다. 회귀계수는 0.2699로 나타났으며, t-통계량은 2.33으로 계산되었다. 이로부터 모형의 예측값과 일별 양매도 전략 사이에 통계적으로 유의한 양의 관계를 확인할 수 있다.

Table 7. The regression result

Measure	Result
R square	0.0087
Coefficient	0.2699
T statistic	2.33

4.2 모형 성능평가

Table 8은 최종 모형의 성능을 평가한 결과이다.

정확도(accuracy)는 모형이 전체 예측 중 얼마나 많이 예측했는지를 나타내는 지표인데, 이 모형은 전체 표본 중 58.2%를 올바르게 예측하였다.

정밀도(precision)는 모형이 긍정(positive)으로 예측한 사례 중 실제로 긍정인 비율이다. 모형의 높은 정밀도는 보수적인 투자전략에 유용하다. 정밀도는 정확한 신호를 의미하기 때문이다. 이 모형은 0.6404의 정밀도를 보였다.

재현율(recall)은 긍정적인 사례 중 얼마나 많은 긍정을 찾아냈는지를 의미한다. 즉, 재현율이 높은 모형은 많은 투자 기회를 포착한다. 이 모형은 0.7828의 재현율을 보였다.

정밀도와 재현율은 서로 상호 보완적인(trade off) 관계에 있어, 투자자의 리스크 허용도, 전략, 그리고 목표에 따라 중요성이 달라질 수 있다[15].

마지막으로, 정밀도와 재현율의 조화 평균을 나타내는 F1 점수는 0.7045로, 양호한 성능을 보였다.

Table 8. Evaluation metrics of result

Metric	Score
Accuracy	0.5820
Precision	0.6404
Recall	0.7828
F1-score	0.7045

4.3 특성 중요도 해석

Fig. 6은 모형의 특성 중요도를 보여주고 있다. 그림을 통해 요일(weekday)이 예측 결과에 가장 큰 영향을 미치고 있음을 알 수 있다. 그리고 VIX 10일 표준편차, 항생지수 3일 표준편차가 중요도 순위에 뒤따랐다.

Global Term/Feature Importances

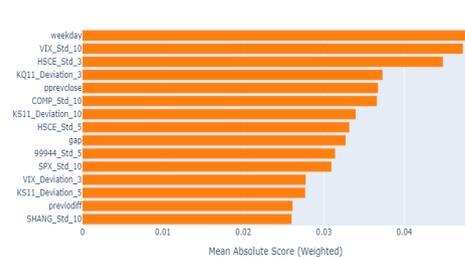


Fig. 6. Feature importances of EBM

Fig. 7, Fig. 8, Fig. 9는 세 변수의 영향력을 그래프 형태로 표현한 것이다. 이를 분석하면, 양매도 전략은 월요일과 양의 관계가 있고, 수요일, 목요일과는 음의 관계가 있다고 해석할 수 있다. 그리고, VIX 10일 표준편차와 항생지수 3일 표준편차가 증가할수록 양매도 전략의 성과에는 부정적인 영향을 미치는 것으로 파악된다. 이 결과는 해외 증시의 변동성이 클 경우, 양매도 전략의 위험성이 높아짐을 의미한다.

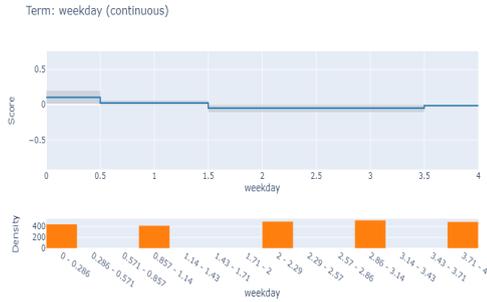


Fig. 7. The score of the weekday to predictions

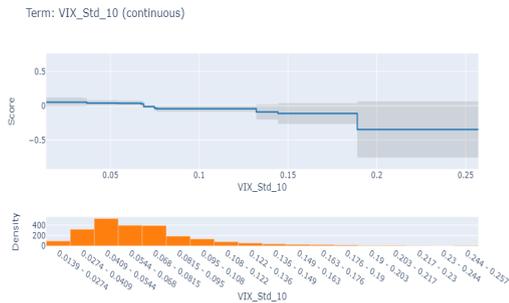


Fig. 8. The score of the VIX_Std_10 to predictions

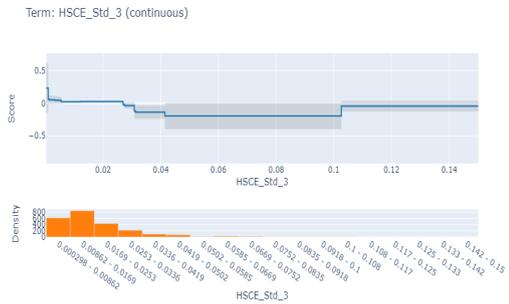


Fig. 9. The score of the HSCE_Std_3 to predictions

4.4 수익성 평가

모형의 예측을 활용하여 매매 시뮬레이션을 진행했다. 이 시뮬레이션은 2016년 3월 23일부터 2018년 12월 28일까지의 테스트 기간을 대상으로 진행했고, 예측값을 확률로 도출하기 위해 ‘predict_proba’ 기능을 사용했다. 또한, 누적수익률은 단리로 산출하였다.

Fig. 10은 모형이 상승 확률을 50% 이상으로 예측한 경우에 매매(이하, S_0.5)를 실행한 수익률 그래프를 보여주고, Fig. 11은 이 상승 확률을 55% 이상으로 예측한 경우(이하, S_0.55)의 매매 결과를 보여주고 있다. 이 두 결과는 매일 양매도를 하는 벤치마크 전략보다 우수한 성과를 보였다. 특히 S_0.55는 큰 하락 구간에 매매를 피하는 경향을 보여주었다.

Table 8에서는 세부적인 성과지표를 확인할 수 있다. S_0.55는 3.53의 샤프비율을 보였고, S_0.5는 2.18의 샤프비율을 기록했다. 이는 벤치마크 전략의 샤프비율인 1.33을 넘어서는 성과이다.

매일 거래하는 벤치마크와는 다르게, S_0.55와 S_0.5는 각각 거래 진입률이 46.56%, 78.03%로 나타났는데, 이는 메타모형을 통해 거래를 필터링하는 방식이 효과적임을 시사한다.

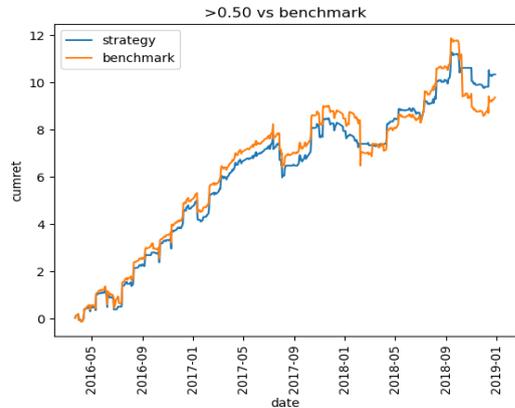


Fig. 10. Equity curve of S_0.5

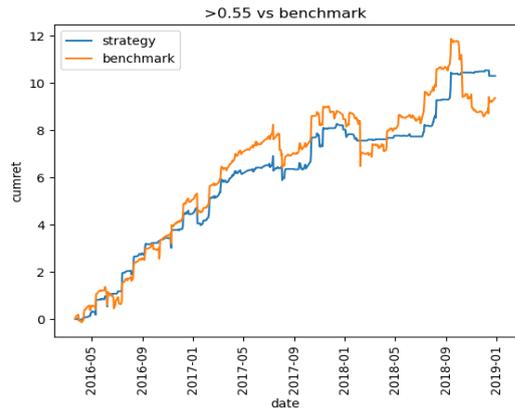


Fig. 11. Equity curve of S_0.55

Table 8. Performance metrics of strategies

Measure	S_0.5	S_0.55	Benchmark
Average	5.43	9.06	3.83
Standard deviation	2.49	2.57	2.88
Shape ratio	2.18	3.53	1.33
Win ratio	65%	68%	65%
Time in market	78.03%	46.56%	100.00%

5. 결론 및 한계점

본 연구는 머신러닝 기법의 하나인 설명가능한 부스팅 머신(EBM)으로 양매도 전략의 진입여부를 결정하는 메타모형을 개발하여, 그 수익성을 분석하였다.

메타모형의 학습을 위해 2004년 1월 5일부터 2018년 12월 28일까지의 외생 변수와 옵션 데이터를 활용하여 95개의 특성을 만들었고, 그중 특성 중요도, T-test, 수동선택으로 22개의 특성을 선택해 양매도 전략과 학습했다.

본 연구의 주요 연구성과는 다음과 같다.

첫째, 메타모형은 0.6404의 정밀도, 0.7828의 재현율, 그리고 0.7045의 F1 점수를 기록했다. 이는 본인의 투자전략이 수익을 낼지, 아니면 손실을 볼지를 약 70%의 확률로 예측할 수 있음을 의미하며, 이는 투자전략 운용에 매우 유용한 정보로 쓰일 수 있다.

둘째, 특성 중요도 분석을 통해 양매도 전략에 영향을 미치는 주요 변수들이 요일, VIX 10일 표준편차, 그리고恒生지수 3일 표준편차임을 확인했다. 이는 EBM의 장점인 설명 가능성(interpretability)을 활용하여 전략에 어떤 변수들이 영향을 미치는지를 파악했다는 점에서 연구의 주요 성과로 간주할 수 있다.

셋째, 제안모형으로 양매도 전략을 진입 여부를 필터링한 결과, 상승을 55%보다 높은 확률로 예측했을 때와 50%를 초과했을 때 매매한 전략의 샤프비율은 각각 3.53과 2.18로, 벤치마크인 1.33을 웃도는 성과를 나타냈다. 이는 메타모형을 통한마켓타임 전략이 효과적인임을 보여준다.

한편, 본 연구는 몇 가지 한계점을 가지고 있다. 첫 번째로, 1차 모형을 단순히 매일 양매도를 진입하는 전략으로 설정한 점이다. 향후 연구에서는 머신러닝이나 기술적 지표를 이용한 매매 전략에 메타모형을 적용하여 전략의 성과를 확인해 볼 필요가 있다.

두 번째로, 최신 데이터를 사용하지 않은 점이 한계점이다. 최근의 코로나와 같은 주요 이벤트가 시장에 큰 영향을 미쳤기 때문에, 최신 자료를 활용한 연구가 필요하다. 더불어 중요한 옵션 데이터인 그리스(greeks)도 수집하여 머신러닝에 활용해 볼 필요가 있다.

References

[1] D. De Giovanni, S. Ortobelli, S. Rachev, "Delta

hedging strategies comparison", *European Journal of Operational Research*, Vol.185, No.3, pp.1615-1631, 2008.

DOI: <https://doi.org/10.1016/j.ejor.2006.08.019>

[2] S.W. Kim, "Profitability of Options Trading Strategy using SVM", *Journal of Convergence for Information Technology*, Vol.10, No.4, pp.46-54, 2020.

DOI: <https://doi.org/10.22156/CS4SMB.2020.10.04.046>

[3] S.W. Kim, "Performance Improvement on Short Volatility Strategy with Asymmetric Spillover Effect and SVM", *Journal of Intelligence and Information Systems*, Vol.26, No.1, pp.119-133, 2020.

DOI: <https://doi.org/10.13088/JIIS.2020.26.1.119>

[4] Y.S. Ra, H.S. Choi, S.W. Kim, "VKOSPI Forecasting and Option Trading Application Using SVM", *Journal of Intelligence and Information Systems*, Vol.22, No.4, pp.177-192, 2016.

DOI: <https://doi.org/10.13088/JIIS.2016.22.4.177>

[5] A. Brunhumer, L. Larcher, P. Seidl, S. Desmettre, J. Kofler, G. Larcher, "Supervised machine learning classification for short straddles on the S&P500", *Risks*, Vol.10, No.12, p.235, 2022.

DOI: <https://doi.org/10.3390/risks10120235>

[6] G.G. Tian, M. Guo, "Interday and intraday volatility: Additional evidence from the Shanghai Stock Exchange", *Review of Quantitative Finance and Accounting*, Vol.28, pp.287-306, 2007.

DOI: <https://doi.org/10.1007/S11156-006-0011-X>

[7] K.R. French, R. Roll, "Stock return variances: The arrival of information and the reaction of traders", *Journal of financial economics*, Vol.17, No.1, pp.5-26, 1986.

[8] S.W. Kim, H.S. Choi, M.G. Bae, "Profitability of intra-day short volatility strategy using volatility risk premium", *Korean Management Science Review*, Vol.27, No.3, pp.33-41, 2010.

[9] M.L. De Prado, *Advances in financial machine learning*, John Wiley & Sons, 2018.

[10] M. Meyer, J.F. Joubert, M. Alfeus, "Meta-Labeling Architecture", *The Journal of Financial Data Science*, Vol.4, No.4, pp.10-24, 2022.

DOI: <https://doi.org/10.3905/jfds.2022.1.108>

[11] H. Nori, S. Jenkins, P. Koch, R. Caruana, "InterpretML: A Unified Framework for Machine Learning Interpretability", 2019.

DOI: <https://doi.org/10.48550/arXiv.1909.09223>

[12] W. Zheng, *TSCV: A Python package for Time Series Cross-Validation*, 2019.

<http://www.zhengwenjie.net/tscv/>

[13] G. Lemaitre, F. Nogueira, C.K. Aridas, "Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning", *The Journal of Machine Learning Research*, Vol.18, No.1, pp.559-563, 2017.

DOI: <https://doi.org/10.48550/arXiv.1609.06570>

- [14] T. Akiba, S. Sano, T. Yanase, T. Ohta, M. Koyama, "Optuna: A next-generation hyperparameter optimization framework", *In Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pp.2623-2631, 2019.
DOI: <https://doi.org/10.1145/3292500.3330701>
- [15] M. Buckland, F. Gey, "The relationship between recall and precision", *Journal of the American society for information science*, Vol.45, No.1, pp.12-19, 1994.
DOI: [https://doi.org/10.1002/\(SICI\)1097-4571\(199401\)45:1<12::AID-AS12>3.0.CO;2-L](https://doi.org/10.1002/(SICI)1097-4571(199401)45:1<12::AID-AS12>3.0.CO;2-L)

김 찬 수(Chansu Kim)

[정회원]



- 2016년 2월 : 인하대학교 토목공학
학과 (산업공학 복수전공) (공학사)
- 2018년 2월 : 국민대학교 비즈니스
스IT전문대학원 (공학석사)
- 2016년 11월 ~ 현재 : 한국펀드평
가 펀드연구원
- 2022년 3월 ~ 현재 : 국민대학교
비즈니스IT전문대학원 박사과정

<관심분야>

알고리즘트레이딩, 계량분석, 머신러닝