

# VIT Feature Extractor 를 활용한 추세 변동성 유사도 기반의 트레이딩 시스템 개발

장주현, 이지훈, 김재윤\*  
순천향대학교

kwack0202@sch.ac.kr, dl123823@gmail.com, \* kimym38@sch.ac.kr

## Development of a Trading System based on Trend Volatility Similarity using VIT Feature Extractor

Jang Juhyeon, Lee Jihun, Kim Jaeyun\*  
\*Soonchunhyang Univ.

### 요 약

본 논문에선 KOSPI200 주가지수의 고빈도 거래데이터를 활용하여 Vision Transformer(VIT)와 추세 다항식 및 볼린저 밴드를 활용해 현재의 시장과 유사한 과거 주가 데이터를 효과적으로 식별하는 트레이딩 시스템을 제안한다. 가격 추세와 시장의 변동성을 정교하게 고려한 거래 전략은 시스템의 생성 신호에 대한 해석 가능성 아래에서 기존의 전략보다 우수한 거래성적을 보였으며 MDD 리스크 관리에서도 안정적인 모습을 확인할 수 있었다.

### I. 서 론

알고리즘 트레이딩 시스템의 개발 초기엔 ARIMA와 같이 수학적 통계모델을 활용하여 주가 시계열의 추세를 효과적으로 포착할 수 있었다.[1] 그러나 높은 시간 의존성으로 인해 시간이 지남에 따라 예측 능력이 낮아지는 문제가 존재해서 실제 시장에서 활용되기 어려운 한계가 있었다. 따라서 최근에는 머신러닝 및 딥러닝과 같이 고도화된 데이터 마이닝 기법을 시스템에 적용하고 있으나 알고리즘이 점점 복잡하고 다층적인 구조가 되면서 시스템 내부의 작동 메커니즘과 의사결정의 원인에 대한 의미론적 해석 가능성이 낮아지고 있다.[2][3] 이로 인한 복잡성은 시스템의 예측 능력과는 대조적으로 매매 의사결정 과정을 이해하기 어렵게 만들어 시스템에 대한 신뢰도를 저하시킬 수 있다. 따라서, 트레이딩 시스템을 주식 시장에서 실제로 응용하기 위해선 생성 신호에 대한 해석 가능성의 균형을 맞추는 것이 필수적이다.

본 논문은 직관적인 해석을 보장하고 효과적인 매매 신호 생성을 위해 추세 다항식을 활용한 캔들스틱 차트 이미지 유사도 기반의 트레이딩 시스템을 제안한다.

### II. 본론

#### 2.1 KOSPI 200 주가지수 선물거래 데이터

본 논문의 제안 트레이딩 시스템을 위한 데이터는 2022년 2월부터 2024년 4월까지 약 2년 3개월에 걸쳐 수집된 KOSPI200 주가지수 데이터를 활용하였다. 해당 데이터는 1분 단위의 짧은 간격으로 수집된 고빈도 거래 데이터이며 상대적으로 긴 시간 간격을 가진 일간 데이터에 비해 더욱 세밀한 시장 움직임을 포착할 수 있다. 본 논문에선 2024년 11월을 기준으로 Train 과 Test 기간을 설정해 6개월의 기간에 걸쳐 제안된 트레이딩 시스템의 매매 성과를 확인하였다.

#### 2.2 캔들스틱 차트 및 볼린저 밴드

캔들스틱 차트는 금융 상품의 가격 변동이 직관적으로 보여지는 시각화 도구이다. 설정된 기간에서 시가와 종가의 가격 변동을 몸통이라 부르며 몸통을 벗어나는 저가와 종가의 변동 범위를 그림자로 부른다. 시가 개시 이후 종가 마감 당시의 가격이 시가보다 상단에 위치할 경우 빨간색, 하단에 위치할 경우 파란색으로 표현된다. 그림.1은 캔들스틱 차트의 예시를 보여준다.

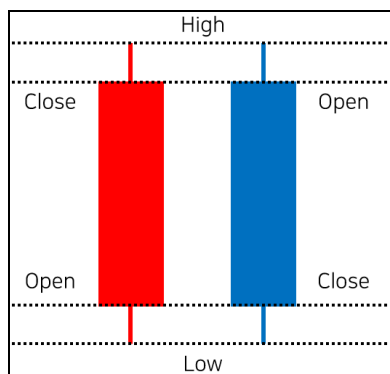


그림.1 캔들스틱 차트 예시

볼린저 밴드는 통계적 개념의 기술적 지표 중 하나로 주가의 단순 이동 평균선과 표준편차를 활용해 주가 변동성이 커지면 표준편차의 값이 증가해 상단과 하단의 밴드가 넓어지고, 변동성이 낮아진다면 반대로 밴드가 좁아진다. 그림.2는 볼린저 밴드 차트를 보여준다

본 논문에서는 차트에 90분의 캔들을 포함했으며 이후 15분에 해당하는 차트를 미래 추세확인 위해 함께 매핑하여 활용하였다. 모든 차트는 하루 단위로 개장 후 90분부터 15분의 간격으로 생성하였으며 일일 평균 21장의 차트 이미지가 생성됐다.



그림.2 볼린저 밴드 차트 예시

### 2.3 추세 다항식 및 Feature Extractor 매매 신호

본 논문에선 90분의 구간과 이후 15분의 구간에 대한 추세 다항식을 생성하여 라벨링을 수행하였다. 각 구간의 추세는 주가 변동이 정규 분포의 특성을 가지는 로그 수익률에 의해 결정된다는 가정 하에, 평균과 표준편차를 활용해 회보를 정의했으며 평균에서 0.43의 표준편차에 해당되는 범위로 설정했다. 이후 구간 별 종가 가격으로 현재 구간의 추세를 추정할 수 있는 1차 다항식을 생성해 회보 범위에서 다항식이 벗어나지 않으면 회보, 상단을 벗어나면 상승, 하단을 벗어나면 하락이 된다.

이미지 생성과 추세 라벨링이 완료된 후에는 VIT모델 구조를 사용해 차트 이미지의 피처를 추출하여 과거 이미지들의 피처와 코사인 유사도 연산 과정을 통해 후보 모집단을 생성하였다. 이후 현재 테스트 구간에 대해서도 동일한 추세 정의 과정을 거쳐 동일한 추세를 보이는 이미지만을 후보 모집단에서 선별했다. 또한, 현재 구간의 볼린저 밴드 영역과 유사한 영역을 보이는 이미지를 정제해 시장 상황이 유사한 차트를 선정하였다. 이러한 단계를 거쳐 구성된 집단에 Hard Voting을 수행해 미래 추세를 종합하여 현재의 추세에 대한 최종 예측을 수행한다. 그림.3은 본 논문에서 제안한 트레이딩 시스템의 프레임워크를 보여준다.



그림.3 제안 트레이딩 시스템 프레임워크

### 2.4 Long & Short Position 진입 및 백테스팅

본 논문은 KOSPI200 주가지수를 활용한 선물거래를 대상으로 트레이딩 전략을 시뮬레이션 하였다. 예측 결과에 따라 Long 또는 Short 포지션을 진입하며, 예측 결과가 현재의 추세와 방향과 일치하면 포지션을 진입하고, 다르다면 No action을 취한다. 포지션 진입 후엔 반대 신호가 예측될 때 포지션을 청산하며, 데이 트레이딩 방식을 적용해 장 마감시에는 무조건 포지션을 청산하도록 한다.

본 논문의 트레이딩 성과를 측정하기 위한 백테스팅 지표는 수익 평균, 손실 평균, payoff ratio, profit factor, MDD를 사용하였다. payoff ratio는 손실 평균 대비 수익 평균을 profit factor는 총 손실 금액 대비 총 수익 금액의 비율을 나타내며, 두 값은 모두 1보다 클 때 수익이 발생함을 의미한다. MDD는 잔고 평가의 고점 대비 최대 낙폭을 의미하는 것으로 투자자가 견뎌야 할 최대 손실 비율을 의미한다. 표1은 각 전략 별 지표를 요약하며, 그림 4는 각 전략 별 누적 수익을 시각화한 것이다.

표1. 제안 트레이딩 시스템 거래 성과

전략	수익평균	손실평균	Payoff ratio	Profit factor	MDD
1_FE	0.735	-0.719	1.022	1.047	-4.001
2_Trend	0.75	-0.583	1.286	1.463	-2.284
3_Bollinger	0.773	-0.595	1.299	1.224	-2.493
BH_L	2.61	-2.37	0.954	0.887	-13.351
BH_S	2.37	-2.261	1.048	1.14	-7.307



그림.4 전략 별 누적 수익 변동 그래프

## III. 결론

본 논문에선 1분 단위로 수집된 고빈도 거래 데이터를 활용해 가격 움직임 및 추세와 시장 변동성이 유사했던 과거 이력을 추출하여 미래의 추세 변동을 예측하였다. 제안 시스템은 알고리즘을 연속적으로 배치해 정교하게 과거 이력을 추출하였으며 payoff ratio와 profit factor를 확인한 결과 비교 전략에 비해 우수한 수익성과 안정적 MDD 리스크 관리를 확인할 수 있었다.

## ACKNOWLEDGMENT

본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업의 연구 결과로 수행되었음. (2021-0-01399) 또한, 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2022R1A2C1092808).

## 참고 문헌

- [1] Pai, Ping-Feng, and Chih-Sheng Lin. "A hybrid ARIMA and support vector machines model in stock price forecasting." *Omega* 33.6 (2005): 497-505.
- [2] Kim, Sang Hoe, et al. "Combining CNN and Grad-CAM for profitability and explainability of investment strategy: Application to the KOSPI 200 futures." *Expert Systems with Applications* 225 (2023): 120086.
- [3] Roostaee, Mohammad Reza, and Ahmad Ali Abin. "Forecasting financial signal for automated trading: An interpretable approach." *Expert Systems with Applications* 211 (2023): 118570.