

## 포트폴리오 최적화와 주가예측을 이용한 투자 모형

Stock Trading Model using Portfolio Optimization and Forecasting Stock Price Movement

---

저자 (Authors)	박강희, 신현정 Kanghee Park, Hyunjung Shin
출처 (Source)	<a href="#">대한산업공학회지 39(6)</a> , 2013.12, 535-545(11 pages) <a href="#">Journal of the Korean Institute of Industrial Engineers 39(6)</a> , 2013.12, 535-545(11 pages)
발행처 (Publisher)	<a href="#">대한산업공학회</a> Korean Institute Of Industrial Engineers
URL	<a href="http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE02311525">http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE02311525</a>
APA Style	박강희, 신현정 (2013). 포트폴리오 최적화와 주가예측을 이용한 투자 모형. 대한산업공학회지, 39(6), 535-545
이용정보 (Accessed)	송실대학교 203.253.***.153 2020/09/29 18:03 (KST)

---

### 저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

### Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

# 포트폴리오 최적화와 주가에측을 이용한 투자 모형

박강희 · 신현정<sup>†</sup>

아주대학교 산업공학과

## Stock Trading Model using Portfolio Optimization and Forecasting Stock Price Movement

Kanghee Park · Hyunjung Shin

Dept. of Industrial Engineering Ajou University

The goal of stock investment is earning high rate or return with stability. To accomplish this goal, using a portfolio that distributes stocks with high rate of return with less variability and a stock price prediction model with high accuracy is required. In this paper, three methods are suggested to require these conditions. First of all, in portfolio re-balance part, Max-Return and Min-Risk (MRMR) model is suggested to earn the largest rate of return with stability. Secondly, Entering/Leaving Rule (E/L) is suggested to upgrade portfolio when particular stock's rate of return is low. Finally, to use outstanding stock price prediction model, a model based on Semi-Supervised Learning (SSL) which was suggested in last research was applied. The suggested methods were validated and applied on stocks which are listed in KOSPI200 from January 2007 to August 2008.

**Keywords:** Portfolio Selection, Markowitz Model, Stock Price Prediction, Semi-Supervised Learning(SSL), Machine Learning

### 1. 서 론

금융투자를 하는 방법은 다양하다. 위험이 적은 곳에 투자하여 적지만 안정적인 수익을 기대하는 투자가 있는가 하면 어떤 투자는 높은 수익을 얻기 위해 상당한 정도의 위험을 감수하면서까지 공격적인 투자를 하는 경우도 있다. 여러 투자 방법이 존재하지만 이들의 궁극적인 목표는 높은 수익률을 얻으면서도 수익률의 변동성이 적은 안정성을 추구하는 것이다. 주식투자에서는, 수익률이 높고 안정적인 종목들로 포트폴리오(Portfolio)를 구성하고, 예측력이 좋은 주가에측 모델을 이용하여 매매(트레이딩)를 하면 이러한 목표에 근접할 수 있다. 이를 위해서, 다양한 포트폴리오 모델 및 주가에측 모델들이 소개되었다.

우선, 대부분의 포트폴리오 모델들은 높은 수익률을 얻기

위해 종목의 투자비율을 효율적으로 배분하는 것을 목표로 한다. 포트폴리오 모델의 대표적인 방법으로는 마코위츠 모델(Markowitz Model)이 있다(Markowitz, 1952). 이후 많은 연구들이 마코위츠 모델을 기반으로 활발히 진행되어 왔다(Brodie, Daubechies, Mol, Giannone, and Loris, 2008; Markowitz, 1952; Seidl, 2012; Steinbach, 2001). Narayan과 Narayan(2010)은 석유 가격의 변동이 베트남 주식시장에 미치는 영향을 분석하여 포트폴리오를 구성하는 방법을 제안하였고, Salem *et al.*(2011)은 S&P 500지수 및 Nikkei 225지수 등 7개국의 주가지수를 바탕으로 포트폴리오를 구성하는 방법을 제안하였다(Narayan and Narayan, 2010; Salem, Shaher, and Khasawneh, 2011). Bekhet과 Matar(2012)는 소비자 지수의 변동에 따른 포트폴리오 구성방법을 제안하였다(Bekhet and Matar, 2012). 금융·경제지수의 영향을 반영하는 포트폴리오 구성방법 뿐만 아니라 종목 간의

본 연구는 아주대학교 연구비지원 및 2013년도 정부의 재원으로 한국연구재단(2010-0007804/2010-0028631/2013R1A1A3010440)의 연구비 지원으로 수행되었음을 사사하며, 이에 감사의 뜻을 표한다(2010-0007804/2010-0028631/2013R1A1A3010440).

<sup>†</sup> 연락처: 신현정 교수, 443-749 경기 수원시 영통구 원천동 아주대학교 산업공학과 Tel: 031-219-2417, Fax: 031-219-1610,

E-mail: shin@ajou.ac.kr

2012년 12월 6일 접수; 2013년 5월 8일 수정본 접수; 2013년 6월 10일 게재 확정.

상관성 및 유사도를 다양한 방법에 적용하여 기술적인 측면을 고려한 포트폴리오 구성방법들도 연구되었다(Andersen, Dahl, and Friberg, 2009; Brodie *et al.*, 2008; Kong and Kim, 2012; Salem *et al.*, 2011).

한편, 주가예측에 대한 모델들도 다양한 방면에서 연구가 진행되어 왔다. Jeantheau(2004)는 Auto-Regressive Conditional Heteroskedasticity(ARCH) 모델을 사용하여 주가를 예측하였고 Amilon(2003)와 Liu *et al.*(2009)은 중국 주식시장에 Skewed-GED Distribution를 이용한 Generalized Auto-Regressive Conditional Heteroskedasticity(GARCH) 모델을 사용하여 주가를 예측하는 방법을 제시하였다(Amilon, 2003; Jeantheau, 2004). Tay and Cao(2001)는 금융시계열 데이터를 Support Vector Machine(SVM)에 도입하는 방법을 제안하였고 Kanas(2003)은 Artificial Neural Network(ANN)모델로 S&P500 지수 예측을 시도하였다(Kanas, 2003; Tay and Cao, 2001). 또한 Yang *et al.*(2001)은 ANN을 이용해 상업은행의 대출리스크(Commercial Bank Loan Risk)관리를 위한 조기경보시스템을 제안하였고 Bekiros and Georgoutsos(2008)는 호재와 악재를 구분하기에 불명확한 뉴스가 NASDAQ 지수에 어떤 영향을 미치는지 ANN 모델을 이용하여 분석하였다(Bekiros and Georgoutsos, 2008; Yang, Li, and Xu, 2001).

그러나 위에서 언급한 포트폴리오 및 주가예측에 대한 연구들에는 다음과 같은 한계가 있다. 첫째, 포트폴리오를 구성하는 종목 간의 투자비율 배분에 관해서는 대부분의 기존연구들이 마코위츠 모델에 기반한다. 즉, 종목 간의 다양성을 추구하면서, 다시 말하자면, 종목 간의 공분산을 최소화 하면서 투자비율을 정하게 되는데, 이는 투자 안정성에 초점이 맞추어진 방법이다. 이러한 방법들에 의하면 수익률의 안정성을 얻을 수는 있겠지만 수익률을 높이는 데는 원천적인 한계가 있다.

둘째, 포트폴리오 모델링은 두 가지 측면, 즉, 리밸런싱(Rebalance)와 업그레이드(Upgrade)로 나누어 볼 수 있다(D. S. Kim and Ryoo, 2007). 포트폴리오 리밸런싱은 운용 중인 포트폴리오의 위험·수익특성에 있어 변동이 발생할 경우 원래의 도하였던 포트폴리오의 특성을 다시 회복하기 위한 수정작업으로서 동일 종목 내에서 투자비율을 재배분하는 활동을 뜻한다 반면 포트폴리오 업그레이드는 새로운 종목을 포트폴리오에 반영하거나 포트폴리오 내의 종목을 제외하는 종목 교체에 해당한다. 그러나 대부분의 기존 연구에서는 마코위츠 모델에 기반한 포트폴리오 리밸런싱 측면에만 초점이 맞추어져 왔다 나머지 다른 측면인 포트폴리오 업그레이드에 대한 내용을 반영하는 연구는 찾아보기 힘들다.

셋째, 기존의 주가예측 모델들은 다음과 같은 한계점이 있다(Park and Shin, 2013). 대부분의 방법론에서는 특정 종목의 미래 주가를 예측하기 위하여 그 종목의 과거 주가에 상당히 의존한다. 이는 미래의 수익률이 과거와 유사할 것이라는 가정을 토대로 하는 것이다. 그러나 주가를 비롯한 각종 금융지표는 국내외적으로 발생하는 여러 요인들과 복잡하게 얽혀있는 금융 네트워크의 일부분으로 볼 수 있다. 즉, 현 시점의 주

가 또는 금융지표는 자신의 과거 데이터보다는 현 시점에서 다른 국내외적인 여러 금융경제지표들의 직·간접적인 영향에 의하여 결정된다고 보는 것이 보다 타당할 것이다(Kim, 2003; Park and Shin, 2013). 기존 연구들은 다른 금융지표나 주가 간의 상호 관계 및 인과관계, 즉, 복잡한 금융 네트워크 구조를 모델에 명시하거나 정형화하는 것에 방법론적으로 한계가 있다.

이러한 한계점들을 극복하기 위해 본 논문에서는 다음의 세 가지 방법론을 제안한다. 첫째, 포트폴리오 모델링의 리밸런싱 측면에서는 구성 종목 간의 분산을 최소화시키는 마코위츠 모델의 안정성을 포함하면서 그와 동시에 수익률 또한 최대화하는 발전된 방법을 제안한다. 제안한 모델은 Max-Return and Min-Risk(MRMR)이라 명명한다. 둘째, 포트폴리오 모델링의 업그레이드 측면에서는 수익률이 저조한 종목을 포트폴리오에서 제외시키고 이 종목과 유사도가 가장 낮은 새로운 종목을 포함시키는 방법을 제안한다. 본 논문에서는 이를 Entering/Leaving Rule(E/L)이라 한다. 셋째, 주가예측모델로는 Park and Shin(2013)이 제안한 Semi-Supervised Learning(SSL) 모델에 기반한 예측모델을 활용한다(Park and Shin, 2013). 이 예측모델은 다양한 금융지표 간의 인과적 복잡성 및 상호연관성을 네트워크로 표현하여 주가를 예측한다.

본 논문은 제2장에서 포트폴리오 구성에 관한 기존 방법인 마코위츠 모델과 주가예측에 활용될 SSL 알고리즘을 설명한다. 제3장에서는 본 논문에서 제안하는 방법들을 소개한다. 포트폴리오 리밸런싱 방법론으로서 Max-Return and Min-Risk를 설명한 후, 포트폴리오 업그레이드 방법론인 Entering and Leaving Rule을 설명한다. 그리고 SSL 기반 주가예측 모델을 설명한다. 제4장에서는 실험 및 결과를 보여준다. 제안한 방법은 2007년 1월부터 2008년 8월의 KOSPI200에 상장된 종목들로 적용 및 검증되었다. 마지막으로 제5장은 결론으로 본 논문을 마무리한다.

## 2. 방법론

### 2.1 마코위츠 모델(Markowitz Model)

마코위츠 포트폴리오 모델은 모든 투자기회 중에서 최적의 리스크, 수익률 조합을 가진 투자기회를 결정하는 이론으로 개별종목의 평균수익률, 분산, 종목간의 과거 자료만 사용하여 분산 투자하는 이론이다. 이 이론은 일정한 수익률 제약 하에 분산이 최소로 하는 종목에 투자하자여 안정적인 포트폴리오를 구성하는 것을 기본 개념을 한다(Markowitz, 1952). 마코위츠 모델을 요약하면 위험의 정도인 종목 간의 분산을 최소화하는 것을 목적함수로 정하고 요구되는 최소 기대수익률을 달성해야 하며, 모든 사용 가능한 금액을 포트폴리오에 투자하고, 공매도가 없다는 세 가지의 제약조건을 가진 비선형계획모델(Nonlinear Programming Model)이다. 모델에 사용되는

변수 및 상수 등을 식(1)과 기호로 정의하면 다음과 같다

$N$  : 포트폴리오에 포함하는 투자 종목의 수  
 $w_i$  : 포트폴리오에서 주식  $i$ 에 투자하는 비율  
 $u_i$  : 주식  $i$ 의 수익률  
 $\sigma_{ij}$  : 주식  $i$ 와  $j$ 의 수익률에 대한 공분산  
 $K$  : 포트폴리오에서 요구하는 최소기대수익률

$$\begin{aligned} \text{Minimize } Z &= \sum_{i,j=1}^N \sigma_{ij} \times w_i \times w_j \\ \text{Subject to } \sum_{i=1}^N u_i \times w_i &\geq K \\ \sum_{i=1}^N w_i &= 1 \\ \forall w_i &\geq 0 \end{aligned} \quad (1)$$

식 (1)의 목적함수는 포트폴리오를 구성하는 종목간의 수익률 공분산을 최소화 한다. 이는 포트폴리오에 투자 위험을 최소화 할 수 있는 종목들의 비율을 높인다. 식 (1)의 첫 번째 제약조건은 기대수익률을  $K$  이상으로 정하여 수익이 낮은 종목의 투자를 피한다. 식 (1)의 두 번째 조건은 사용 가능한 투자를 모두 투자하는 제약식이다. 마코위츠 모델은 이 조건들을 만족시키면서 포트폴리오의 위험을 최소화 하는 최적의 투자비율을 구하게 된다. 본 모형은 비선형계획법 모형 중에서 목적함수가 한계 체감(Decreasing Marginal Return)을 보이는 이차계획(Quadratic Programming)으로 최적해를 구한다(Hillier and Hillier, 2008).

## 2.2 그래프 기반 준지도 학습 (Graph-based Semi-Supervised Learning, SSL)

그래프 기반 준지도 학습(Graph-based semi-supervised learning, SSL)에서는 그래프로 표현된 유사도를 이용하여 레이블이 없는 데이터의 레이블을 예측한다(Shin, Hill, Lisewski, and Park, 2010; Shin, Lisewski, and Lichtarge, 2007; Zhou, Bousquet,

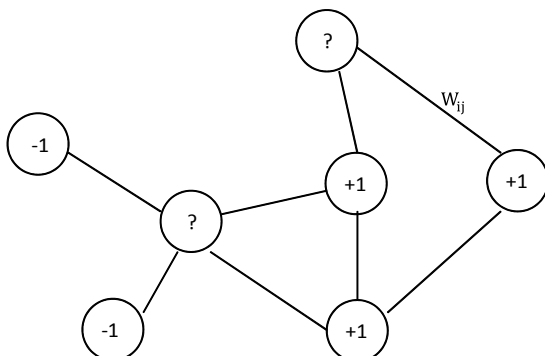


Figure 1. Graph-based semi-supervised learning(SSL)

Lal, Weston, and Schölkopf, 2004). <Figure 1>은 SSL의 기본 모형을 보여준다. SSL 방법은 전체 데이터  $n(n=1+u)$ 개 중에서  $l$ 개의 레이블이 있는 데이터  $\{(X_1, Y_1), \dots, (X_l, Y_l)\}$ 와  $u$ 개의 레이블이 정해지지 않은 데이터를 사용한다. 데이터의 레이블은 레이블이 있을 경우 -1 또는 1로 설정하며( $Y_l \in \{-1, 1\}$ ,  $l=1, \dots, l$ ) 그렇지 않은 경우 0으로 설정한다( $Y_u \in \{0\}$ ,  $u=1+l, \dots, l+u$ ).

데이터들은 그래프의 노드들로 표현되고 유사도가 높다고 판단되는 노드들 사이에는 엣지가 형성된다.  $i$ 노드와  $j$ 노드 사이의 엣지의 연결강도는 유사도가 증가할수록 커지는데 다음의 식 (2)에 의해 결정된다. 예를 들어,  $i$ 노드와  $j$ 노드의 유사도가 높다면 식 (2)에 따라  $w_{ij}$ 값이 커지게 되고,  $i$ 노드와  $j$ 노드의 유사도가 낮다면  $w_{ij}$ 값이 작아지게 된다. 여기서  $\sigma^2$ 은 분산을 의미한다.

$$w_{ij} = \begin{cases} \exp\left(\frac{(x_i - x_j)^T (x_i - x_j)}{\sigma^2}\right) & \text{if } i \sim j \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

각 지표들간의 상관성 및 인과관계를 나타내는 것은 유사행렬  $W$ 를 사용한다. 유사행렬  $W$ 를 만드는 방법으로는 주로 K-Nearest Neighbor(KNN) 또는 일정 반경 내에서의 유클리드 거리를 활용한 방법이 주로 사용된다( $\|x_i - x_j\|^2 < r$ ). SSL 알고리즘은 식 (3)의 이차목적함수를 최소화하여 레이블이 없는 노드들에 대한 예측값  $f$ 를 출력한다(Belkin, Matveeva, and Niyogi, 2003; Belkin and Niyogi, 2004; Zhou, Bousquet, Lal, Weston, and Schölkopf, 2004).

$$\text{Min}_f (f - y)^T (f - y) + \mu f^T L f \quad (3)$$

여기서  $y$ 는 목표값으로서  $y = (y_1, \dots, y_l, 0, \dots, 0)^T$ 로 설정되며 출력값(예측값)은  $f = (f_1, \dots, f_l, f_{l+1}, \dots, f_{l+u})^T$ 로 표현된다.  $L$ 은 Graph Laplacian Matrix 이며  $D = \text{diag}(d_i)$ ,  $L = D - W$ 로 정의된다. 식 (3)의 두 가지 학습조건인 앞부분(loss function)을 (a), 뒷부분(smoothness function)을 (b)라고 하면, (a) 출력값  $f$ 는 레이블이 된 노드에서는 노드의 목표 값  $y$ 와 비슷해야 하고, (b) 연결된 두 노드  $i$ 와  $j$ 의 출력값은 서로 비슷해지는 제약을 만족시키는 해를 산출한다. 식 (3)의 파라미터  $\mu$ 는 이러한 두 가지 조건이 학습에 미치는 영향을 조절하는 학습 파라미터로서 사용자에게 의해 정의된다. 식 (3)에서 최소화하는 조건을 만족시키는 해를 구하기 위해 미분을 하면 다음의 식 (4)를 유도하여 출력값  $f$ 를 계산할 수 있다.

$$f = (I + \mu L)^{-1} y. \quad (4)$$

$I$ 는 단위행렬을 의미한다.

### 3. 제안 방법론

우수한 수익을 내는 주식투자를 위해서는 수익률이 높고 변동이 안정적인 종목들로 구성된 포트폴리오(Portfolio)와 예측력이 정확한 주가예측 모델이 필요하다. 본 논문에서는 이것을 이루기 위하여 다음의 세 가지 방법을 제안한다. 첫째, 포트폴리오 리밸런스 측면에서 안정성에만 기반을 두어 설계된 마코위츠 모델의 한계를 극복하기 위해 안정성을 포함하면서 그와 동시에 최대의 수익률을 얻을 수 있는 Max-Return and Min-Risk(MRMR) 모델을 제안한다. 둘째, 포트폴리오 업그레이드 측면에서 수익률이 낮은 종목을 제거하고 수익률이 높은 종목을 포함시켜, 포트폴리오 구성종목을 효율적으로 관리할 수 있는 Entering/Leaving Rule을 제안한다. 셋째, 주가예측 모델 측면에서는 정확도가 높은 SSL기반 예측모델을 이용하여 Buy and Sell Strategy를 시도한다. <Figure 2>는 위 절차를 간략하게 나타낸다.

다음 각 절에서는 본 연구목적에 적합하게 제안하는 Max-Return and Min-Risk(MRMR) 모델, Entering/Leaving Rule, SSL기반 Buy and Sell Strategy에 대하여 간략히 소개한다.

#### 3.1 포트폴리오 리밸런스 방법론 : Max-Return and Min-Risk Model(MRMR)

마코위츠 모델은 리스크로 간주하는 종목 간의 공분산을 최소화하여 종목 투자비율을 산출한다. 공분산이 최소화가 되면 리스크는 안정적이나 최상의 수익률을 얻기는 힘들어진따라서 본 논문에서 제안하는 Max-Return and Min-Risk(MRMR) 모델은 마코위츠 모델의 한계를 다음과 같이 개선한다. 포트폴리오 구성종목 간의 공분산을 최소화 시키는 동시에 수익

률을 최대화하는 방법으로 목적함수가 설계된다. 즉, 마코위츠 모델의 목적함수인 종목 간 공분산 최소화 대신 MRMR 모델의 목적함수는 종목 간 공분산을 분모로 설정하고 포트폴리오 기대수익률을 분자로 설정하고 이를 최대화한다. 이는 포트폴리오의 리스크 감소에만 초점이 맞추어져 있던 마코위츠 모델과 달리 리스크 감소와 동시에 수익률 향상을 추구할 수 있는 개선된 방법이다. MRMR 모델의 목적함수와 제약조건은 다음 식 (5)과 같다.

- $N$  : 포트폴리오에 포함하는 투자 종목의 수
- $w_i$  : 포트폴리오에서 주식  $i$ 에 투자하는 비율
- $u_i$  : 주식  $i$ 의 수익률
- $\sigma_{ij}$  : 주식  $i$ 와  $j$ 의 수익률에 대한 공분산
- $K$  : 포트폴리오에서 요구하는 최소기대수익률
- $k$  : 개별종목 최소기대수익률
- $R$  : KOSPI 등락율
- $c$  : 최소기대수익률 조정상수

$$\text{Maximize } Z = \frac{\sum_{i=1}^N u_i \times w_i}{\sum_{i=1}^N \sum_{j=1}^N \sigma_{ij} \times w_i \times w_j} \quad (5)$$

$$\text{Subject to } \sum_{i=1}^N u_i \times w_i \geq K$$

$$u_i \geq k \left( k = \frac{R}{C} \right)$$

$$\sum_{i=1}^N w_i = 1$$

$$\forall w_i \geq 0$$

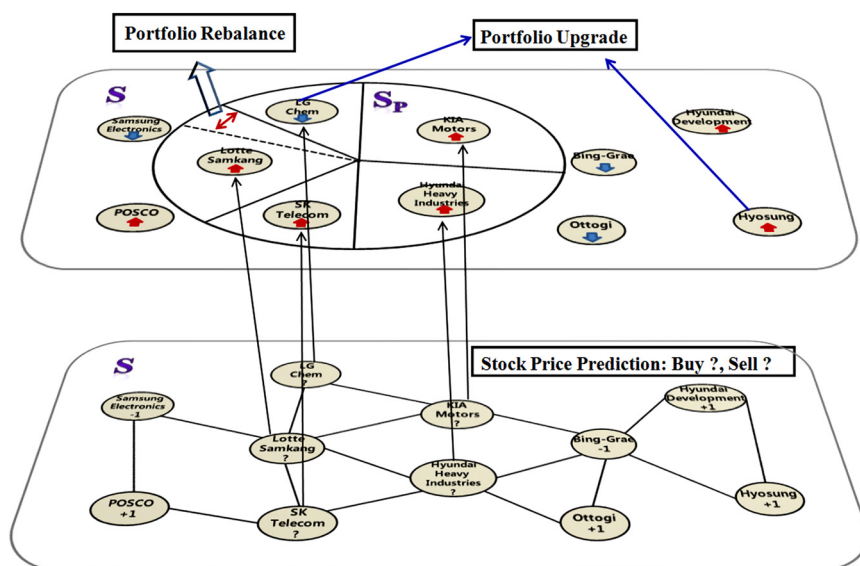


Figure 2. Structure of proposed three methods: Max-Return and Min-Risk model (MRMR), Entering/Leaving rule (E/L), Buy and Sell Strategy based on semi-supervised learning

제약조건 역시 마코위츠 모델의 최소 기대수익률 부분에서 한 가지를 추가하였다. 마코위츠 모델에서는 포트폴리오 전체의 최소기대수익률( $K$ ) 이상의 제약조건만 주어졌으나 MRMR 모델에서는 개별종목 최소기대수익률( $k$ ) 조건을 추가했다. 이는 수익률이 낮은 종목의 투자비율이 늘어나는 문제점이 있을 경우, 이를 조정해 주는 제약조건이다. 개별종목의 최소기대수익률은 모델생성시점의 KOSPI 등락율과 비교하여 일정 값 이상으로 정의한다. 이 값은 조정계수  $c$ 에 따라 탄력적으로 조정 가능하게 설계하였다. MRMR 모델은 포트폴리오 내의 종목 간 투자비율만을 조정하므로 포트폴리오 업그레이드 측면은 고려할 수 없다. 다음 절에서는 이점을 보완하기 위해 포트폴리오를 업그레이드하는 방법인 Entering/Leaving Rule을 제안한다.

### 3.2 포트폴리오 업그레이드 방법론 : Entering/Leaving Rule(E/L)

MRMR 모델은 포트폴리오 수익률이 하락할 경우 포트폴리오 리밸런싱 즉, 구성종목의 비율만 조정한다. 그러나 포트폴리오를 구성하는 대부분의 종목이 하락할 경우에는 수익률 하락을 피하기 어렵다. 이를 해결하기 위해서 구성종목을 바꾸는 포

트폴리오 업그레이드가 필요하다. 본 절에서는 포트폴리오 업그레이드의 방법론인 Entering/Leaving Rule(E/L)을 제안한다.

E/L은 다음과 같다. MRMR 모델의 초기 포트폴리오 구성종목은 과거수익률이 우수한 종목으로 구성한다. 그리고 E/L은 제 3.1절 식 (5)의 개별종목 최소기대수익률( $k$ )에 부합하지 않는 종목을 Leaving시킨다. 그리고 전체종목 중 그 종목과 가장 유사도가 낮은 종목을 선택하여 포트폴리오 구성종목에 포함시킨다. 유사도는 제 2.2절의 식 (2)를 이용하여 계산한다. 여기에서 포트폴리오에 Entering하는 종목의 과거 수익률은  $k$ 에 상관없이 선택한다. 즉, 새로 선택하는 종목의 과거수익률은 최소 기대수익률의 조건과 관계없이 유사도에 따라서만 선택한다. 왜냐하면, Entering 종목의 과거수익률은 유사도를 구할 때 일부 부분 영향을 미치므로, 만약 Entering종목의 과거수익률이 식 (5)의 최소기대 수익률( $k$ )에 미치지 못하더라도 유사도가 가장 낮은 해당종목을 선택한다. 예를 들어 E/L을 도식화하면 다음과 <Figure 3>과 같다.

전체종목 집합을  $S$ 라 정의하고 포트폴리오 구성집합을  $S_p$ 라고 정의하면, 포트폴리오  $S_p$ 의 구성종목 중 LG chem의 수익률이 개별종목 최소기대수익률보다 낮을 경우 이 종목을 Leaving시킨다. 그리고 LG chem과 가장 유사도가 낮은 종목인 Hyosung을 포트폴리오  $S_p$ 의 구성종목에 새로 Entering시킨다. 이를 요약하면 다음 <Figure 4>와 같다.

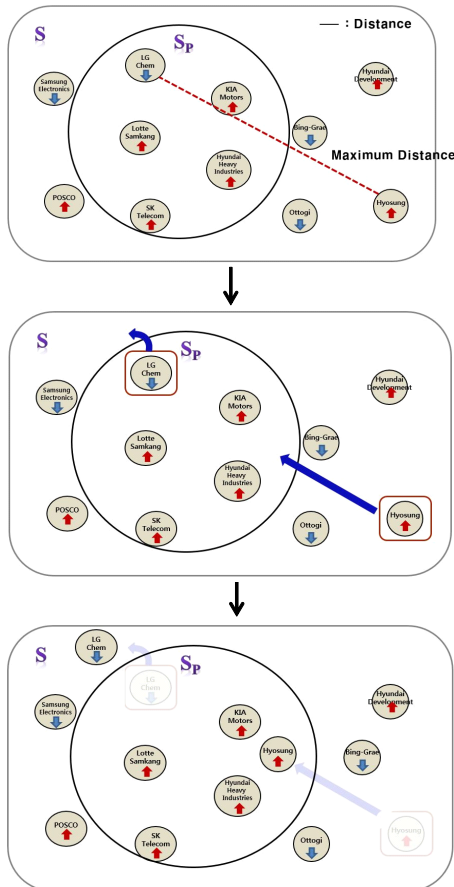


Figure 3. Changes of stocks which composite portfolio based on Entering/Leaving rule

#### Definition

- $u_i$  :  $i$ 번째 종목의 실제수익률
- $S$  : 전체종목그룹
- $S_p$  : MRMR 모델을 구성한 종목그룹
- $s_i$  :  $i$ 번째 개별종목
- $k_i$  :  $i$ 번째 종목의 최소기대수익률
- $s_i^q$  :  $s_i$ 와 가장 유사하지 않는(dissimilar) 종목
- $x$  : 전체종목그룹의 개별종목

#### Procedure

```

For  $i = 1 : N$ 
  if  $u_i < k_i$ 
     $S_p = S_p \setminus \{s_i\}$ 
     $s_i^q = \max_x \text{dissimilarity}(s_i, x)$ ,
    dissimilarity는 식 (2)를 이용
     $S_p = S_p \cup s_i^q$ 
  end
end
    
```

Figure 4. The definition of Entering/Leaving Rule

E/L에 의해 교체되는 구성종목을 바탕으로 MRMR 모델이 새로 업그레이드되면 그에 따른 투자비율도 <Figure 5>와 같이 바뀌게 된다.



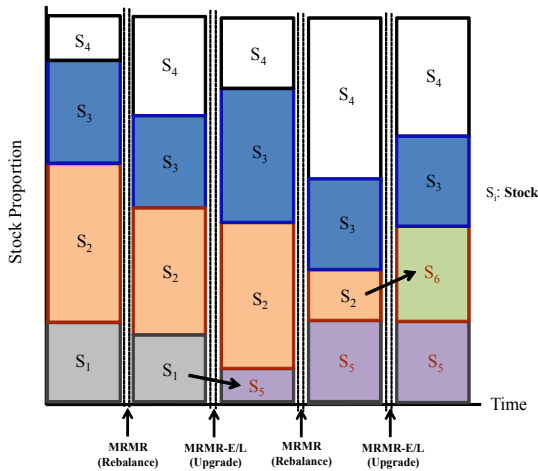


Figure 5. Changes of stocks and ratio by MRMR-E/L

### 3.3 주가예측모델 : SSL기반 Buy and Sell Strategy

Buy and Sell strategy를 이용하여 트레이딩을 하기 위해서는 주가예측 모델이 필요하다. 본 연구에서는 시계열 분석에 적용된 SSL을 이용하여 Buy and Sell Strategy를 시도한다. 시계열 분석에 적용한 SSL은 노드가 주가 등의 금융지표들로 표시되고 노드간의 엣지는 금융지표간의 유사도로 표시된다(Zhou *et al.*, 2004). 노드들에는 레이블이 주어지는데 시계열 데이터로 이루어진 각 지표들은 연속형의 값을 가지므로 이분형 레이블을 가질 수 있게 변형을 시도한다. 이분형의 레이블로 설정하기 위해 식 (6)의 방법으로 해당지표의 등락을 이용하였다

$$y = \text{sign}(X_t - \text{MA}_5(X_t)). \quad (6)$$

t시점의 데이터를  $X_t$ 라 할 때 레이블의 의미는 주가와 5일 이동평균값을 비교하여 주가가 높을 때 +1, 이동평균값이 높을 때 -1로 설정된다. 노드(금융지표)에 레이블이 포함된 그래프 기반 SSL은 <Figure 6>과 같은 기본 모형으로 표현된다.

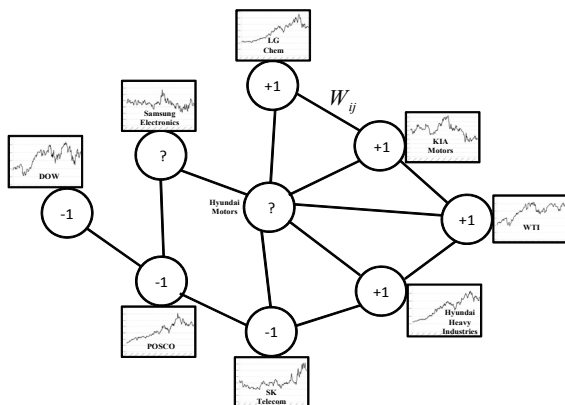


Figure 6. SSL representation for time series prediction

SSL은 유사도 매트릭스(W)의 구성방법에 따라 모델의 예측 성능을 좌우한다. 그런데 시계열 데이터는 고유한 잡음(oscillatory noise)을 가지고 있다. 예측성능이 높은 유사도 매트릭스를 구성하기 위해서는 시간의 연관성을 보존하면서 잡음을 줄이는 방법이 필요하다. 따라서 Kim(2006) 및 Park *et al.*(2011)에서 제안한 시계열 데이터를 기술적 지표(Technical Indicators)로 변환시키는 방법을 이용한다

기술적 지표는 금융분석에서 자주 사용하는 방법인데 기술적 지표화의 장점은 시계열 데이터가 고유하게 갖고 있는 잡음(oscillatory noise)을 제거하고 그 밑에 존재하는 큰 구조 즉 추세 및 구조적 변동요인들을 이끌어 내는 역할을 한다. 주가와 금융지표들은 변수 특성상 시계열 데이터로 이루어져 있다. 이 시계열 데이터는 시간에 따른 연관관계를 형성하고 식(7)과 같이 표현된다.

$$X_t = \{x_1, x_2, \dots, x_j, \dots, x_t\} \quad (7)$$

여기서  $x_j$ 는 t를 전체 시간으로 할 때 i번째 시점의 데이터를 의미한다.  $X_t$ 가 시계열 데이터인 특성상 SSL에 직접 적용하기에는 몇 가지 문제점이 존재한다. <Figure 6>에서 보면 그래프의 전체 노드들은 식 (7)과 같이 각 노드마다의 고유 시계열 데이터를 갖고 있다.

예를 들어 <Figure 6>에서 POSCO 노드도 시계열 데이터  $X_t^{\text{KIA-Motors}}$ 가 있고 KIA-Motors노드 또한 고유의 시계열 데이터  $X_t^{\text{KIA-Motors}}$ 가 있다. 문제는, 이들 간의 유사도를 두 시계

Table 1. The Definition of Technical Indicators(Tis)

	TIs	Meaning
$s_1$	$\text{MA}_p(X_t) = \frac{1}{p}(x_t) + \frac{p-1}{p}\text{MA}_p(X_{t-1})$	p-moving average (exponential smoothing)
$s_2$	$\text{BIAS}_p(X_t) = \frac{x_t - \text{MA}_p(X_t)}{\text{MA}_p(X_t)}$	The change rate of $x_t$ relative to $\text{MA}_p(X_t)$
$s_3$	$\text{OSC}_{p,q}(X_t) = \frac{\text{MA}_p(X_t) - \text{MA}_q(X_t)}{\text{MA}_p(X_t)}$	The change rate of $\text{MA}_q(X_t)$ relative to $\text{MA}_p(X_t)$
$s_4$	$\text{ROC}_p(X_t) = \frac{x_t - x_{t-p}}{x_t}$	The relative rate of change for $X_t$ between p consecutive time points
$s_5$	$K_t^p = \frac{X_t - \text{Min}_{i=t-p-1}^t(x_i)}{\text{Max}_{i=t-p-1}^t(x_i) - \text{Min}_{i=t-p-1}^t(x_i)}$	Standardization of $x_t$
$s_6$	$D_t^p = \text{MA}_3(K_t^p)$	3-Moving Average of $K_t^p$
$s_7$	$\text{RSI}_t^p = \frac{\sum_{i=t-p-1}^t ( x_1 - x_{1-1} )}{\sum_{i=t-p-1}^t ( x_1 - x_{1-1} )}$	The relative strength index.

열 데이터들로부터 직접적으로 도출하기가 어렵다는데 있다. 따라서, 각 시계열 데이터들을 구조상의 요인들 즉 추세 변동을 반영하는 기술적 지표로 변환하고 이로부터 두 노드 POSCO와 KIA-Motors간의 유사도를 측정한다. 다음 <Table 1>은 본 연구에서 시계열 데이터를 변환시키는데 사용한 7가지 기술적 지표들을 정리한 것이다. 계산된 노드간 7개의 기술적 지표는 식 (2)을 이용하여 유사도 매트릭스가 구성된다.

기술적 지표로 계산된 유사도 매트릭스와 전체종목의 주어진 레이블로, 포트폴리오에 구성된 종목의 레이블을 예측하고 이 값에 따라 Buy and Sell Strategy를 시도한다. 예측값  $f_t$ 가 ' $f_t > 0$ '이면 예측시점에서 매수를 시도하고 ' $f_t < 0$ '이면 매도를 시도한다. 이를 도식화하면 <Figure 7>과 같이 표현된다.

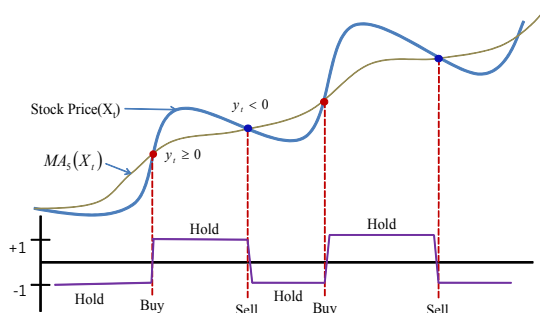


Figure 7. Buy and Sell Strategy: trading rule using SSL predicted value

## 4. 실험

### 4.1 데이터

본 실험에서 사용한 데이터는 2007년 1월부터 2008년 8월까지 총 403개의 일별 주가(daily data point)로 이루어져 있다. 포트폴리오 모델에 사용한 데이터는 KOSPI200에 상장된 200개 회사들이다. <부록 1>에 상장회사들의 리스트가 제시되어 있으며 이들은 10개의 산업별로 구성되어 있다. Buy and Sell Strategy를 하기 위해 SSL을 이용한 주가예측 모델에 사용한 데이터는 KOSPI200에 상장된 200개 회사를 포함할 뿐만 아니라 주요 글로벌 경제지수인 DOW, NASDAQ, NIKKEI, HSI, SSE, TSEC, FTSE, DAX, CAC, BSE\_SENSEX, IBOVESPA, AORD, KOSPI, Exchange\_Rate(KRW-USD), WTI, CD(Certificate of Deposit)이 이용하였다. 독립변수와 종속변수는 다음과 같다. 독립변수는 포트폴리오에 구성되지 않은 종목 및 금융지수 206개다. 종속변수는 포트폴리오에 구성되는 10개 종목으로 이루어진다.

### 4.2 모델 비교 방법

실험에서는 본 연구에서 제안한 MRMR-E/L 모델과 E/L을 적용시키지 않은 MRMR 모델, E/L을 적용한 마코위츠 모델

마코위츠 모델을 비교한다. 2007년 1월부터 2007년 5월까지의 96개의 일별 데이터를 학습 및 검증기간(training and validation period)으로 설정하였고, 2007년 6월부터 2008년 8월까지의 307개의 일별 데이터에 대한 성능을 비교하였다(Liu et al., 2009).

포트폴리오 최초 모델에 구성되는 종목은 <부록 1>의 10개 분야에서 학습 및 검증기간 동안 가장 수익률이 높은 종목을 각각 한 종목씩 선택하여 총 10개로 구성하였다. 포트폴리오 업그레이드 주기는 테스트기간 내에 20개씩 시점으로 설정하였고, E/L을 하기 위해 종목 간의 유사도는 SSL 알고리즘의 유사도 매트릭스를 이용하였다(Park and Shin, 2013; Ryoo, 2007). E/L을 적용하지 않은 모델들은 위와 동일한 주기로 포트폴리오 리밸런싱을 실시하였다.

주가예측 모델인 SSL은 Rolling Forecast 방법을 사용하여 예측하였다(M, W, and K, 2000). Rolling Forecast 방법이란 1에서  $t$ 시점까지 데이터를 이용하여  $t+1$ 시점을 예측하고,  $t+2$ 시점을 예측할 때는 2시점부터  $t+1$ 시점까지를 학습기간으로 사용한다. SSL 모델에 설정되어야 하는 파라미터로는  $k$ 와  $\mu$ 가 있으며, 이들은 각각 식 (2)의 이웃노드 개수와 식 (3)의 loss-smoothness tradeoff를 의미한다. 실험에 사용된 파라미터의 값들은  $\{k, \mu\} \in \{2, 3, 4, 5\} \times \{0.01, 0.1, 0.3, 0.5, 0.7, 1, 10, 100\}$  범위의 값들 중 검증셋에 대하여 최적의 조합으로 선정하였다.

모델간의 수익률 비교방법은 ROI를 사용하였다(Barber, Lehavy, McNichols, and Trueman, 2001). ROI는 투자대비 얻는 수익을 계산하는 방법으로서 다음 <Figure 8>과 같은 의미를 나타내고 다음 식 (8)을 따른다.

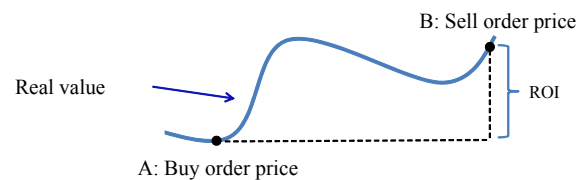


Figure 8. The Definition of return on investment(ROI)

$$ROI = \frac{B-A}{A} \times 100(\%) \quad (8)$$

### 4.3 결과

#### 4.3.1 수익률 비교

수익률을 비교하기 위하여 제안하는 모델인 MRMR-E/L과 MRMR, 마코위츠-E/L 그리고 마코위츠 모델의 수익률에 대한 성과를 제시한다. <Figure 9>는 테스트기간에서 E/L에 따라 포트폴리오 구성이 변하는 종목들을 표기했다.

2007년 6월부터 타임포인트가 20시점씩 지날 때마다 E/L을 시도하였고, 각 시점 별로 E/L이 일어날 때마다 포트폴리오 구성에 1개~6개 종목의 변화가 나타났다. 2008년 6월 25일 이후



부터 8월까지(테스트기간의 끝) 포트폴리오에 속한 종목들이 E/L의 기준에 모두 부합하여 변화가 없었다 E/L을 적용하지 않는 모델은 기존의 10종목을 바꾸지 않고 20개의 시점마다 종목 간의 비율만 조정하였다 <Table 2>는 각 모델들은 테스트 기간의 20시점 구간 내에서 SSL을 이용한 예측값에 따라

Buy and Sell Strategy의 방법에 따른 ROI 값을 비교한 것이다 (Barber *et al.*, 2001).

매매수수료는 온라인 최저 수수료인 0.015%로 가정하여 16개의 테스트 구간의 매매 결과 KOSPI가 평균 -11.31%(±7.36%) 하락하는 동안 동일기간에 MRMR-E/L 모델은 +13.72%

2007-06-01	2007-07-27	2007-08-27	2007-09-27	2007-10-26	2007-11-23	2007-12-24
LG화학	LG화학	LG화학	LG화학	대한제분	대한제분	금호타이어
LG생명과학	LG생명과학	LG생명과학	기아차	롯데제과	롯데제과	롯데제과
삼성화재	LG디스플레이	LG디스플레이	LG디스플레이	LG디스플레이	남양유업	LG디스플레이
포스코	포스코	포스코	태광산업	KT	KT	롯데칠성음료
현대차	남양유업	남양유업	현대차	현대차	현대차	경방
삼성전자	삼성전자	롯데삼강	우리금융지주	대경기계기술	대우건설	한국포이름
SK텔레콤	SK텔레콤	SK텔레콤	SK텔레콤	SK텔레콤	SK텔레콤	우리금융지주
KT&G	KT&G	KT&G	롯데칠성음료	롯데칠성음료	삼성SDI	삼성SDI
CJ	CJ	CJ	대우건설	대우건설	크라운제과	크라운제과
대림산업	대림산업	대림산업	대림산업	고려제강	기아차	제일기획

2008-01-24	2008-02-26	2008-03-25	2008-04-23	2008-05-27	2008-06-25
태광산업	KB국민은행	KB국민은행	KB국민은행	태광산업	태광산업
기아차	기아차	기아차	영풍	영풍	영풍
LG디스플레이	LG디스플레이	동원F&B	동원F&B	동원F&B	동원F&B
신한금융지주	태광산업	태광산업	하이닉스	하이닉스	롯데제과
현대차	현대차	현대차	현대차	현대시멘트	기아차
한국포이름	한국포이름	현대상선	대한유화공업	대한유화공업	KT
일성신약	한화	한화	하이트맥주	우리금융지주	일성신약
삼성SDI	삼성SDI	삼성SDI	롯데삼강	롯데삼강	대우건설
크라운제과	하이닉스	하이닉스	롯데제과	롯데제과	두산인프라
LG전자	LG전자	LG전자	LG전자	대한제분	대한제분

Figure 9. Changes of composing stocks during test-period using E/L

Table 2. Each models ROI in 16 sections during test-period

Period	MRMR-E/L	MRMR	Markowitz-E/L	Markowitz	KOSPI
T1 2007.06.01~2007.06.27	5.86%	3.57%	2.51%	3.60%	4.27%
T2 2007.06.28~2007.07.26	11.30%	11.25%	-5.35%	11.28%	12.11%
T3 2007.07.27~2007.08.24	-3.48%	-4.33%	-13.32%	-4.35%	-4.89%
T4 2007.08.27~2007.09.21	3.09%	7.85%	-3.34%	7.85%	6.43%
T5 2007.09.27~2007.10.25	-0.56%	-1.02%	-3.27%	-1.37%	1.59%
T6 2007.10.26~2007.11.22	2.20%	-2.22%	-5.10%	-2.22%	-11.29%
T7 2007.11.23~2007.12.21	-1.11%	7.59%	-2.61%	7.59%	5.98%
T8 2007.12.24~2008.01.23	0.99%	-9.53%	-1.08%	-9.30%	-15.16%
T9 2008.01.24~2008.02.25	3.90%	-14.08%	-14.41%	-14.08%	2.77%
T10 2008.02.26~2008.03.24	1.16%	-3.75%	-9.02%	-3.68%	-3.16%
T11 2008.03.25~2008.04.22	11.10%	8.34%	3.91%	8.34%	6.75%
T12 2008.04.23~2008.05.26	-8.83%	5.00%	-15.28%	5.06%	0.00%
T13 2008.05.27~2008.06.24	-0.87%	-1.98%	-6.30%	-1.96%	-6.30%
T14 2008.06.25~2008.07.22	-13.60%	-7.83%	-10.24%	-7.83%	-9.09%
T15 2008.07.23~2008.08.20	3.27%	4.79%	-3.68%	0.00%	-3.21%
T16 2008.08.21~2008.08.29	-0.56%	-1.32%	-0.52%	-2.76%	-2.51%
ROI (±Std)	13.72% (±6.27%)	2.31% (±7.17%)	-87.12% (±5.72%)	-3.83% (±7.07%)	-11.31% (±7.36%)

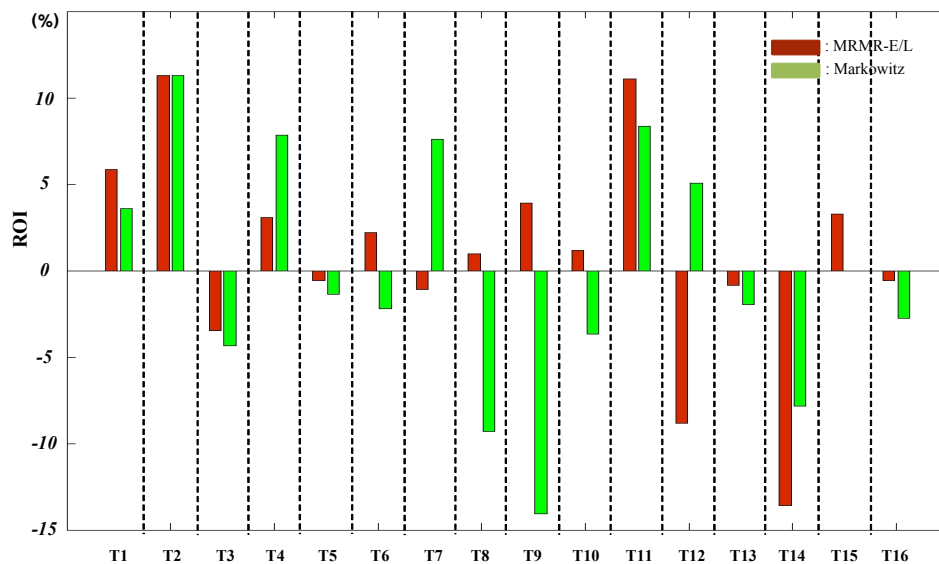


Figure 10. ROI comparison between models in 16 sections : MRMR-E/L model vs. Markowitz model

( $\pm 6.27\%$ )의 수익률을 나타낸 반면 MRMR 모델, 마코위츠-E/L 모델 그리고 마코위츠 모델은 각각  $+2.31\%$ ( $\pm 7.17\%$ ),  $-87.12\%$ ( $\pm 5.72\%$ ) 그리고  $-3.83\%$ ( $\pm 7.36\%$ )의 평균수익률을 나타내었다. MRMR-E/L 모델의 수익률은 상대적으로 비교 모델보다 최소  $+10.41\%$ 에서 최대  $+100.84\%$ 의 우수함이 보였다. 동일기간 KOSPI의 등락률보다 좋은 수익률을 나타낸 MRMR-E/L 모델, MRMR 모델 그리고 마코위츠 모델의 안정성을 서로 비교해도 MRMR-E/L 모델의 표준편차가 가장 낮음을 보여 상대적으로 다른 모델보다 더 안정적임을 알 수 있다. <Figure 10>은 제안한 방법들의 성능을 비교하기 위해 MRMR-E/L 모델의 구간별 수익률과 마코위츠 모델의 구간별 수익률을 도식화한 것이다. MRMR-E/L 모델이 16개의 전체 테스트 구간 중 12구간에서 마코위츠 모델보다 높게 나타나 제안한 방법이 우수함을 보였다. 제안한 방법의 추가적인 검증을 위해 <부록 1>의 분야를 5개로 재분류하고, 제 4.2절과 같은 방법으로 5종목을 뽑아 <부록 2>에 모델 별 수익률을 비교하였다. 5종목일 때 역시 제안한 방법이 가장 우수한 수익률을 나타내었다.

## 5. 결론

본 논문은 포트폴리오 리밸런싱 측면과 포트폴리오 업그레이드 측면에서 MRMR 모델과 E/L을 이용하여 리스크를 최소화하는 동시에 수익률을 증가시키는 새로운 포트폴리오를 구성하는 방법을 제안하였다. 또한 주가예측모델 측면에서 예측력이 좋은 SSL 알고리즘을 이용하여 트레이딩을 시도하였다. 본 연구에서 제안하는 방법은 다음과 같은 이점이 있다. 첫째, 안정성을 중점으로 설계된 마코위츠 모델의 약점을 보완하여 안정성을 포함하면서 최대의 수익률을 구할 수 있다. 즉, 기술적으로 설명하면 마코위츠 모델은 종목 간의 공분산을 최소화하

여 유사도가 떨어지는 종목에 투자비율을 증가시키는 방법이 었다. 그러나 본 논문에서 제안하는 MRMR 모델은 종목 간의 분산을 최소화 시킬 뿐만 아니라 수익률이 좋은 종목의 투자비율을 최적으로 조정하여 수익률을 높일 수 있는 장점을 가진다. 둘째, 포트폴리오를 업그레이드할 때 수익률이 좋지 않은 종목을 포트폴리오 구성에서 제외하고, 제외한 종목과 가장 유사도가 낮은 새로운 종목을 포트폴리오 구성에 추가하여 수익률을 향상시켰다. 이 방법은 수익률이 낮은 종목에 투자되어 수익률이 감소하는 현상을 차단하고 수익률이 높은 종목들을 포트폴리오에 새로 구성하는 효과가 있어 수익률 향상에 도움을 주었다. 마지막으로 주가예측 모델은 이전연구에서 예측력이 우수하다고 알려진 SSL 알고리즘을 이용하여 좋은 수익률을 얻을 수 있었다. 이를 요약하면 포트폴리오 구성방법은 유사도가 낮은 종목을 중심으로 수익률의 최적화를 이루었고, 주가예측은 유사도가 비슷한 종목을 통해 예측하였다. 이러한 이점들의 결합으로 동일 테스트 기간에서 마코위츠 모델 등의 비교 모델보다 ROI가 상대적으로  $10.41\% \sim 100.84\%$ 의 더 우수한 수익률을 얻었다. 현재 제안된 연구는 투자예산이 충분하다고 가정하고 모델을 제안하였다. 그러나 현실에서는 예산의 한계가 존재할 수 있다. 이를 고려하여 향후에 예산의 한계에 따라 투자비율을 조정할 수 있는 모델로 개선된다면 좀 더 현실에서 잘 이용할 수 있는 연구가 될 것이다.

## 참고문헌

- Amilon, H. (2003), GARCH estimation and discrete stock prices : an application to low-priced Australian stocks, *Economics Letters*, **81**(2), 215-222. doi : 10.1016/S0165-1765(03)00172-1.
- Andersen, E. D., Dahl, J., and Friberg, H. A. (2009), Markowitz port-

- folio optimization using MOSEK, *MOSEK Technical Report*, **2**, 1-30.
- Barber, B., Lehavy, R., McNichols, M., and Trueman, B. (2001), Can Investors Profit from the Prophets? Security Analyst Recommendations and Stock Returns, *The Journal of Finance*, **1**, 531-563.
- Bekhet, H. A. and Matar, A. (2012), Risk-Adjusted Performance : A two-mdel Approach Application in Amman Stock Exchange, *International Journal of Business and Social Science*, **3**(7), 34-45.
- Bekiros, S. and Georgoutsos, D. (2008), Direction-of-Change Forecasting using a Volatility-Based Recurrent Neural Network, *Journal of Forecasting*, **27**, 407-417.
- Belkin, M., Matveeva, I., and Niyogi, P. (2003), Regression and Regularization on Large, In : *Shawe-Taylor, J., Singer, Y. (eds.) COLT 2004. LNCS (LNAI)*, 3120, 624-638.
- Belkin, M. and Niyogi, P. (2004), Semi-Supervised Learning on Riemannian Manifolds, *Machine Learning*, **56**, 209-239.
- Brodie, J., Daubechies, I., Mol, C. D., Giannone, D., and Loris, I. (2008), Sparse and stable Markowitz portfolios *Working paper series*: European Central Bank.
- Hillier, F. S. and Hillier, M. S. (2008), *Introduction to Management Science*.
- Jeantheau, T. (2004), A link between complete models with stochastic volatility and ARCH models, *Finance Stochast*, **8**, 111-131. doi : 10.1007/s00780-003-0103-6.
- Kanas, A. (2003), Non-linear forecasts of stock returns, *Journal of Forecasting*, **22**(4), 299-315. doi : 10.1002/for.858.
- Kim, D. S. and Ryoo, H. S. (2007), Portfolio Management Using Statistical Process Control Chart, *IE Interfaces*, **20**(2), 94-102.
- Kim, K.-J. (2003), Financial time series forecasting using supportn vector machines, *Neurocomputing*, **55**, 307-319.
- Kim, K.-J. (2006), Artificial neural networks with evolutionary instance selection for financial forecasting, *Expert Systems with Applications*, **30**, 519-526.
- Kong, M. and Kim, J. (2012), The Study on Volatility in Stock Market, *Korean Journal of Business Administration*, **25**(2), 953-969.
- Liu, Q., Sung, A. H., Chen, Z., Liu, J., Huang, X., and Deng, Y. (2009), Feature Selection and Classification of MAQC-II Breast Cancer and Multiple Myeloma Microarray Gene Expression Data, *MAQC-II Gene Expression*, **4**(12), 1-24.
- M, O. C., W, R., and K, G. (2000), Does updating judgmental forecasts improve forecast accuracy?, *International Journal of Forecasting*, **16**(1), 101-109.
- Markowitz, H. (1952), Portfolio Selection, *The Journal of Finance*, **7**(1), 77-91.
- Narayan, P. K. and Narayan, S. (2010), Modelling the impact of oil prices on Vietnam's stock prices, *Applied Energy*.
- Park, K., Hou, T., and Shin, H. (2011), Oil Price Forecasting Based on Machine Learning Techniques, *Journal of the Korean Institute of Industrial Engineers*, **37**(1), 64-73.
- Park, K. and Shin, H. (2013), Stock Price Prediction based on a Complex Interrelation Network of Economic Factors, *Engineering Applications of Artificial Intelligence*, **26**(5-6), 1550-1561.
- Ryoo, H. (2007), A compact mean-variance-skewness model for large-scale portfolio optimization and its application to the NYSE market, *Journal of the Operational Research Society*, **58**, 505-515. doi : 10.1057/palgrave.jors.2602168.
- Salem, R., Shaher, T. A., and Khasawneh, O. (2011), International Portfolio Diversification Benefits for Middle Eastern Investors, *Journal of Money, Investment and Banking*, **22**, 22-31.
- Seidl, I. (2012), Markowitz versus regime switching : an empirical approach, *The review of finance and banking*, **4**(1), 033-043.
- Shin, H., Hill, N. J., Lisewski, A. M., and Park, J.-S. (2010), Graph sharpening, *Expert Systems with Applications*, **37**(12), 7870-7879. doi : 10.1016/j.eswa.2010.04.050.
- Shin, H., Lisewski, A. M., and Lichtarge, O. (2007), Graph sharpening plus graph integration: a synergy that improves protein functional classification, *Bioinformatics*, **23**, 3217-3224. doi : 10.1093/bioinformatics/btm511.
- Steinbach, M. C. (2001), Markowitz Revisited : Mean-Variance Models in Financial Portfolio Analysis, *Society for Industrial and Applied Mathematics*, **43**(1), 31-85.
- Tay, F. E. H. and Cao, L. (2001), Application of support vector machines in financial time series forecasting, *Omega*, **29**(4), 309-317. doi: 10.1016/S0305-0483(01)00026-3.
- Yang, B., Li, L. X., and Xu, J. (2001), An early warning system for loan risk assessment using artificial neural networks, *Knowledge-Based Systems*, **14**(5-6), 303-306.
- Zhou, D., Bousquet, O., Lal, T. N., Weston, J., and Schölkopf, B. (2004), Learning with Local and Global Consistency, *Advances in Neural Information Processing Systems*, **16**, 321-328.

## &lt;Appendix 1&gt; List of KOSPI200

화학	LG 화학, S-Oil, 호남석유화학, LG생활건강, 아모레퍼시픽, 한국타이어, 금호석유화학, KCC, 효성, 넥센타이어, SKC, 태광산업, 금호타이어, 카프로, 후성, 대한유화공업(주), 남해화학, 유니드, 케이퍼케미칼, 을촌화학, SK케미칼, 삼성정밀화학, 한국화인케미칼, 한화석유화학, 동양제철화학, 동화약품공업, 휴켄스
의약품	LG생명과학, 유한양행, 동아제약, 녹십자, 일양약품, 부광약품, 대웅제약, 종근당, 일성신약, 중외제약, 일동제약, 한미약품, 광동제약, 영진약품공업
금융업	삼성화재해상보험, 신한금융지주회사, 하나금융지주보통부, 우리금융지주, 중소기업은행, 삼성증권, 대우증권, 미래에셋증권, 우리투자증권, 현대증권, 코리안리재보험, 대구은행, 부산은행, 한국투자금융지주
철광금속+비철금속	포스코, 현대제철, 고려아연, 현대하이스코, 세아베스틸, 영풍, 동국제강, 풍산, 고려제강, 세아제강, 삼광유리공업, 쌍용양회공업, 한일시멘트, 조선내화, 한국철강, 유니온스틸, 아세아시멘트, 대한제강
기계장비+운수장비	현대차, 현대모비스, 기아차, 현대중공업, 두산중공업, 삼성중공업, 대우조선해양, 두산인프라코어, 한라공조, 현대미포조선, 현대엘리베이터, STX조선, 한진중공업, 쌍용자동차, S&T중공업, STX엔진, 동양기전, S&T대우, 동부제강, 현대시멘트, 비엔지스틸
전기전자	삼성전자, 하이닉스반도체, LG전자, 삼성SDI, 삼성전기, 삼성테크윈, LS산전, 대덕전자, 대한전선, 신도리코, 한국단자공업, 대덕GDS, 일진전기, 삼성전자공업, 금호전기, 한미반도체, S&TC
서비스업+통신업	SK텔레콤, LG, KT, 삼성엔지니어링, 엔씨소프트, SK, 강원랜드, 두산, 웅진코웨이, 동원시스템즈, 제일기획, 에스원, SBS, 코오롱, 대교, 세림제지, LG디스플레이, 현대오트넷
제조업+운수창고업	KT&G, 현대상선, 대한항공, 한화, 동부한농, 대상, 대한제분, 경방, 한진해운, 한국유리공업, 태평양, 금호산업(주), 한솔LCD, LS전선, 한국기술산업, 무림페이퍼, 한국전기초자, 대경기계기술
음식료업	CJ, 오리온, 삼성물산, 현대건설, 제일모직, 농심, 삼양사, 빙그레, 남양유업, 롯데삼강, 오투기, LG상사, 현대산업개발, 롯데제과, 롯데칠성음료, 크라운제과, 하이트맥주, 동양메이저, 한국제지, 삼양제넥스
기타	대림산업, 롯데쇼핑, GS건설, 현대백화점, 대우건설, 한국가스공사, 한국전력공사, 대우인터내셔널, SK네트웍스, 신세계, 한섬, 한솔제지, 더베이직하우스, LG Fashion Corporation, 명성, 신성이엔지, 셀러, FnC코오롱, 성신양화, 애경유화, 새한, 한국코트렐, 나산, Fursys 한국카본, 한국포리올, 케이씨텍, KTF (KT합병), 동원F&B, GS홀딩스, 세원셀론텍

## &lt;Appendix 2&gt; Comparing each models ROI in 16 sections during test-period while selecting 5 stocks

초기 선택 종목 (분야)	LG화학 (화학+의약품)	SK텔레콤 (금융업+서비스업 +통신업)	포스코 (철광금속+비철금 속+기타)	삼성전자 (전기전자+ 제조업+운수창고업)	현대차 (기계장비+운수장비 + 음식료업)
모델	MRMR-E/L	MRMR	Markowitz-E/L	Markowitz	KOSPI
전체기간 (2007.06.01~2008.08.29) 수익률 (±표준편차)	<b>15.38%</b> <b>(±10.03%)</b>	4.64% (±15.74%)	-15.31% (±8.86%)	3.78% (±20.50%)	-11.31% (±7.36%)