

SVM을 이용한 VKOSPI 일 중 변화 예측과 실제 옵션 매매에의 적용

라윤선

국민대학교 비즈니스IT전문대학원
(90dbstjs@naver.com)

최흥식

국민대학교 비즈니스IT전문대학원
(hschoi@kookmin.ac.kr)

김선웅

국민대학교 비즈니스IT전문대학원
(swkim@kookmin.ac.kr)

기계학습(Machine Learning)은 인공 지능의 한 분야로, 데이터를 이용하여 기계를 학습시켜 기계 스스로가 데이터 분석 및 예측을 하게 만드는 것과 관련한 컴퓨터 과학의 한 영역을 일컫는다. 그중에서 SVM(Support Vector Machines)은 주로 분류와 회귀 분석을 목적으로 사용되는 모델이다. 어느 두 집단에 속한 데이터들에 대한 정보를 얻었을 때, SVM 모델은 주어진 데이터 집합을 바탕으로 하여 새로운 데이터가 어느 집단에 속할지를 판단해준다. 최근 들어서 많은 금융전문가는 기계학습과 막대한 데이터가 존재하는 금융 분야와의 접목 가능성을 보며 기계학습에 집중하고 있다. 그러면서 각 금융사는 고도화된 알고리즘과 빅데이터를 통해 여러 금융업무 수행이 가능한 로봇(Robot)과 투자전문가(Advisor)의 합성어인 로보어드바이저(Robo-Advisor) 서비스를 발 빠르게 제공하기 시작했다. 따라서 현재의 금융 동향을 고려하여 본 연구에서는 기계학습 방법의 하나인 SVM을 활용하여 매매성과를 올리는 방법에 대해 제안하고자 한다. SVM을 통한 예측대상은 한국형 변동성지수인 VKOSPI이다. VKOSPI는 금융파생상품의 한 종류인 옵션의 가격에 영향을 미친다. VKOSPI는 흔히 말하는 변동성과 같고 VKOSPI 값은 옵션의 종류와 관계없이 옵션 가격과 정비례하는 특성이 있다. 그러므로 VKOSPI의 정확한 예측은 옵션 매매에서의 수익을 낼 수 있는 중요한 요소 중 하나이다. 지금까지 기계학습을 기반으로 한 VKOSPI의 예측을 다룬 연구는 없었다. 본 연구에서는 SVM을 통해 일 중의 VKOSPI를 예측하였고, 예측 내용을 바탕으로 옵션 매매에 대한 적용 가능 여부를 실험하였으며 실제로 향상된 매매 성과가 나타남을 증명하였다.

주제어 : 기계학습, SVM, VKOSPI, 옵션 매매

논문접수일 : 2016년 11월 1일 논문수정일 : 2016년 12월 13일 게재확정일 : 2016년 12월 16일
원고유형 : 일반논문 교신저자 : 최흥식

1. 서론

금융시장에는 주식, 선물, 옵션 등과 같은 여러 종류의 투자 가능한 상품이 있다. 이러한 상품을 매매하는 투자자의 관심사는 미래에 해당 상품 가격의 상승 또는 하락에 대한 방향성(Directional Movement) 예측이다. 미래 가격의 방향성을 정확하게 예측할 수 있다면 투자자들

은 높은 투자 수익을 얻을 수 있다. 방향성 예측에 대한 전통적인 방법으로는 기업의 본질가치(Fundamental Value)를 분석하는 기본적 분석(Fundamental Analysis)(Sharma, 2016)과 가격, 거래량 등의 정보를 이용하여 가격의 방향성을 분석하는 기술적 분석(Technical Analysis) 방법이 대표적이다. 최근에는 글로벌 금융시장의 복잡성이 확대됨에 따라 기계학습 방법이 각광을 받

기 시작하였다. 이와 관련하여 기계학습의 인공 신경망(Artificial Neural Network) 모델을 바탕으로 주가를 예측한 사례가 있었다(Jeong and Yun, 1998; Siddiqui and Abdullah, 2015). 또한, SVM 모델을 바탕으로 여러 개의 시스템 트레이딩 전략 중 수익을 가져다줄 것이라고 판단되는 전략만을 골라 전략 포트폴리오를 구성하는 방법을 제시한 연구도 있었다(Park et al., 2014).

그동안 금융시장에서 가격의 방향성 예측에 대한 많은 연구가 있었다(Choi et al., 2011). 그러나 대부분은 가격의 무작위성(Random Walk)을 기각하지 못하고 있다. 그만큼 금융시장에서 주가의 방향성 예측은 쉽지 않다. 주가의 방향성 예측과 비교해서 가격 변동의 크기를 측정하는 변동성(Volatility)에 대한 예측은 변동성의 군집(Clustering) 특성으로 인해 방향성보다는 상대적으로 예측 확률이 높은 편이다. 주식시장의 변동성을 측정하는 대응치로는 역사적 변동성(Historical Volatility), 내재변동성(Implied Volatility), 그리고 변동성지수(Volatility Index) 등이 사용되고 있다. 많은 연구 결과들은 변동성 지수의 미래 변동성에 대한 예측력이 우수함을 밝히고 있다(Choi and Lee, 2010). 한국거래소는 한국형 변동성지수인 VKOSPI¹⁾를 투자자들에게 발표하고 있다. VKOSPI는 KOSPI200을 기초자산으로 하는 옵션 가격을 토대로 KOSPI200 지수의 향후 변동성을 측정한 지수로서, 미국에서 S&P500 옵션 가격을 토대로 산출하는 VIX지수와 비슷하다.

VKOSPI는 시황변동의 위험을 감지하는 투자 지표로서 변동성 위험을 관리할 수 있는 투자수단의 역할을 할 것으로 기대된다(Kim, 2010). 그

러므로 가까운 미래의 VKOSPI 예측에 대한 연구는 상당히 의미 있는 분석 대상이다. VKOSPI는 투자자들의 심리가 많이 반영되기 때문에 주식시장 내의 투자자들의 심리상태가 반영된 여러 지표를 통해 어느 정도 예측이 가능하며 또한 세계 금융시장 동조화로 인한 변동성 전이 효과로 인해 외국 장의 변동성 변화를 토대로 VKOSPI 추측이 가능하다(Jeong et al., 2012; Rajhans and Jain, 2015). 특히 전일 미국의 VIX 지표의 움직임에 큰 영향을 받기 때문에 전일 VIX 지표를 이용하는 것도 VKOSPI 예측에 도움이 될 수 있다(Moon, 2016).

투자자들은 VKOSPI 예측 결과를 여러 금융상품 매매에 이용할 수 있다. ELS²⁾(Equity Linked Securities, 주가연계증권)에 투자하는 경우 변동성 변화에 따라 원금 손실이 발생할 가능성이 달라진다(Lim and Choi, 2015). 따라서 ELS 투자자들은 변동성을 예의주시해야 한다.

무엇보다 변동성과 가장 연관이 깊은 투자상품은 옵션(options)이다(Cho, 2015). 옵션 가격은 VKOSPI에 영향을 많이 받는다. 옵션의 특징으로는 만기, 행사가격에 따라 종류가 상당히 많다는 것이다. 투자자는 여러 종류의 옵션을 이용하여 옵션 포지션을 구축함으로써 다양한 손익구조를 설계할 수 있고, 따라서 VKOSPI의 방향 예측이 가능하다면 이를 이용하여 수익을 낼 수 있는 포지션 구축 또한 가능하다. 구체적으로 주식시장 내 거래가 이루어지는 동안의 VKOSPI 변화, 즉 일 중 변화를 알 수 있다면 수익을 기록하는데 도움이 된다. 본 연구에서는 기계학습 방법 중 시계열자료에서 우수한 예측력을 보이는 SVM 기법을 활용하여 VKOSPI의 일 중 예측도

- 1) 한국거래소(KRX)가 2009년 4월 13일부터 국내주식시장에 맞게 고안해 낸 아시아가 최초의 변동성지수.
- 2) 개별 주식의 가격이나 주가지수에 연계되어 투자수익이 결정되는 유가증권이다.

형을 제안하고, 실제 옵션 시장에서의 가격 자료를 이용하여 변동성 매도전략을 실증분석하여 연구 성과를 분석하고자 한다.

본 논문의 구성은 다음과 같다. 2장에서는 기계학습 방법의 하나인 SVM 이론에 대해 설명한다. 3장에서는 실제로 SVM 패키지를 이용하여 일 중 특정 시간대의 VKOSPI를 예측하고 실제 결과와의 비교를 통해 예측성과를 살펴본다. 4장에서는 예측된 일 중 VKOSPI를 옵션 매매에 어떻게 적용하여 매매성과를 높일지에 대한 아이디어를 소개하고, 최종적으로 과거 시뮬레이션 결과를 토대로 향상된 매매 성과를 설명한다. 마지막 5장에서는 연구에 대한 결론과 향후 연구 방향에 대해 제시한다.

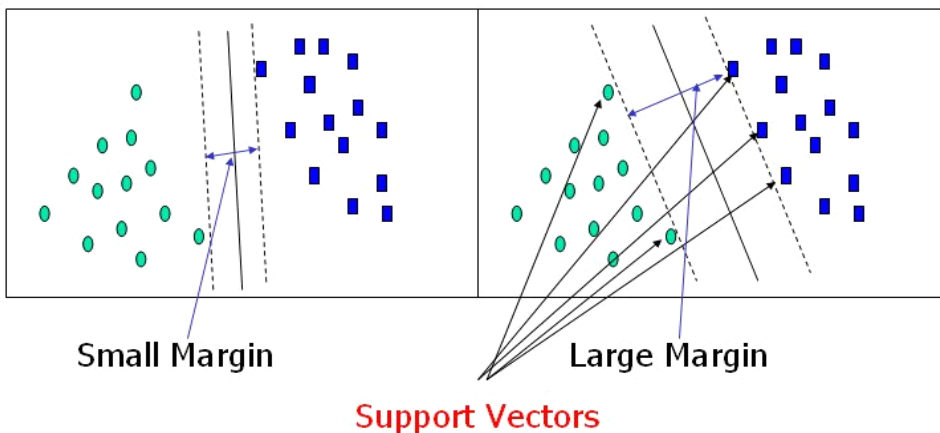
2. Support Vector Machines 이론

SVM은 기계학습의 분야 중 하나로 주로 분류를 위해 사용되며 서로 다른 두 집단에 속한 데이터들에 관한 정보를 얻었을 때, SVM 알고리즘

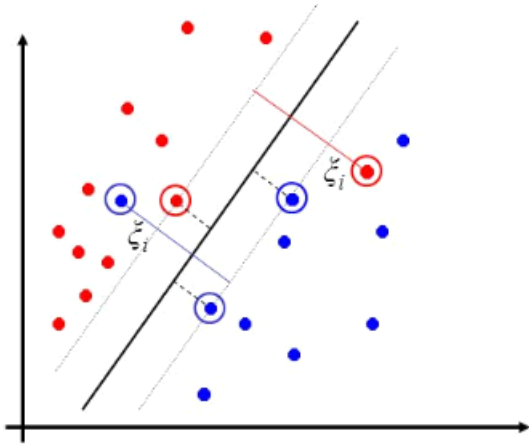
은 주어진 데이터 집합을 바탕으로 하여 새로운 데이터가 어느 집단에 속할지를 판단해준다. SVM 알고리즘은 두 집단의 데이터를 적당한 벡터 공간에 사상(Mapping)시키고 난 후 적당한 분류 기준을 공간상에 표현해 준다. 이후 새로운 데이터가 다시 공간 내로 진입했을 때 SVM 모델은 새로운 데이터의 위치에 따라 그 데이터가 어느 집단에 속하는지 판단해준다. SVM은 기계학습 방법 중에서도 인공지능과 더불어 예측력이 우수하다고 알려져 있기에 많은 분야에서 SVM 모델을 이용한 연구를 진행하고 있다(Park and Hansen, 2012; Li et al., 2008).

2.1 선형 SVM

기본적인 SVM 모델은 선형 SVM으로 <Figure 1>과 같이 오분류(Miss-Classification)없이 선형식(Linear Equation)으로 데이터가 완전히 분리되는 경우이다. 이때 각 데이터 집단의 경계에 있는 점들을 서포트 벡터(Support Vectors)라고 하며, 서로 다른 집단의 서포트 벡터 사이의 직선



<Figure 1> Maximum-margin hyperplane



<Figure 2> Linear SVM with miss-classification

거리를 마진(Margin)이라고 한다. SVM 모델에서는 마진이 최대화되도록 선형식이 결정되고 이를 최대 마진 초평면(Maximum-Margin Hyperplane)이라고 한다. 최대 마진 초평면은 새로운 데이터를 분류해주는 기준이 된다.

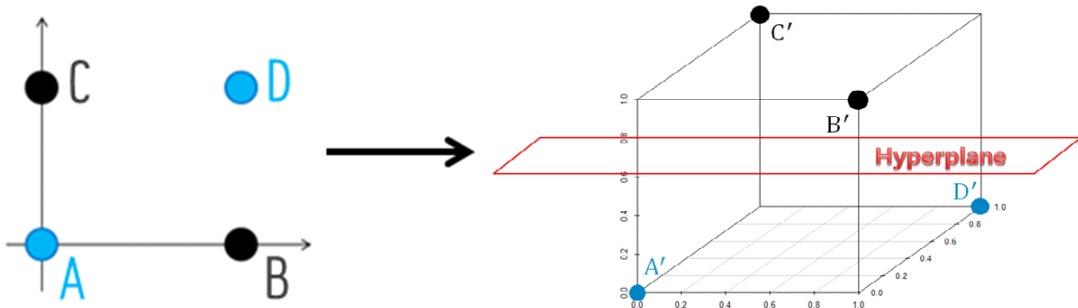
<Figure 2>와 같이 오분류 없이 데이터를 완전히 선형 분리할 수 없는 경우에는 오분류를 인정하되 오분류 정도를 최소화하는 방향으로 선형 SVM이 결정된다. 이때 SVM 모델은 ‘최대 마진’, ‘오분류 최소화’ 두 가지의 목적을 적절히 고려하며 선형식을 결정한다. 사용자는 모델 적

용 시 C라는 매개변수의 설정을 통해 두 가지의 목적의 비중을 조정할 수 있다(Oh, 2008; Flach, 2012).

2.2 비선형 SVM

1992년 최대 마진 초평면을 구하는 문제에 커널 트릭(Kernel Trick)을 적용하는 비선형 분류가 제안되었다(Boser et al., 1992). 이를 일컬어 비선형 SVM이라고 하는데 핵심은 ‘고차원 공간으로의 데이터 이동’이다.

<Figure 3>은 기존 공간(2차원)에서 선형분리가 되지 않던 데이터들이 고차원 공간(3차원)으로 이동함으로써 선형분리가 되는 것을 보여준다. 데이터가 이동된 고차원 공간에서의 초평면은 기존 공간에서는 선형이 아닌 비선형이므로 ‘비선형 SVM’ 용어가 사용된다. <Figure 3>과 같이 때로는 고차원 공간으로의 데이터 이동을 통한 비선형 SVM이 분류 성과를 더 높이기도 한다. 하지만 실제로는 고차원 공간으로의 사상을 가능하게 해주는 함수를 매번 찾을 수 없다. 다만 고차원 공간으로의 이동을 위해 커널 함수(Kernel Function)를 이용한다. 커널 함수는 비선형 SVM을 가능케 해주며 또한 커널 함수의 종



<Figure 3> Move data to higher dimensional space

류 선택은 분류 성과를 결정지어주는 중요한 요소가 될 수 있다(Kim and Ahn, 2010; Tay and Cao, 2002). 대부분의 프로그래밍 언어에서 SVM 구현이 가능하고, 그중 R에서는 ‘e1071’이라는 패키지를 통해 SVM 모델링을 할 수 있다. <Table 1>은 ‘e1071’ 패키지의 커널 함수 4가지이다.

<Table 1> Four kinds of kernel functions

Kernel functions	Linear
	Polynomial
	Radial basis
	Sigmoid

3. SVM 모델에 의한 일 중 VKOSPI 방향 예측

SVM 모델을 통해 특정 태스크를 구현하고 싶은 사용자들을 위해 R, Python, Java, C/C++, MATLAB과 같은 프로그래밍 언어에서는 SVM 패키지를 무료로 제공한다. 사용자는 프로그래밍 상에 자신의 데이터를 불러와서 SVM을 직접 모델링 할 수 있다. 본 연구에서는 R 프로그래밍을 통해 SVM 모델링을 구현한다(Lantz, 2013).

예측 대상은 VKOSPI의 일 중 움직임이며 구체적으로 09:30분 ~ 15:00분 사이의 변화이다. 모델링을 위해 09:00분 ~ 09:30분까지 이후 일 중 VKOSPI 변화에 영향을 줄 수 있다고 판단한 몇 가지의 Input 변수의 값들을 수집하고, 해당일 09:30분 ~ 15:00분 동안의 VKOSPI의 실제 변동 값을 구하여 (+)부호 혹은 (-)부호에 맞춰 각각 ‘Up’, ‘Down’으로 새로 이름 지어(Labeling)

Output 값을 할당했다. 그 후 Input 변수와 Output 변수 간의 인과관계를 기반으로 SVM 모델링을 실행하였다.

3.1 Input Data

VKOSPI는 주식시장 내의 투자자 심리의 불안 함 정도를 객관적으로 수치화한 값이므로 투자자들의 심리를 나타내주는 몇 가지의 변수를 선정하여 <Table 2>에 나타내었다.

<Table 2> Input data

Variable name	KOSPI Return
	KOSPI Volatility
	Net foreign buying amount of KOSPI
	Net foreign-invested futures quantity
	Individual call option net buying quantity
	Individual put option net buying quantity
	Open interest
	VIX close on the day before close

Input Data는 총 8개로 이루어져 있다. 이 값들은 09:30분에 수집이 완료된다. 8개 중 ‘KOSPI 수익률’, ‘KOSPI 변동 폭’, ‘KOSPI 외국인 순매수 금액’, ‘외국인 선물 순매수 수량’, ‘개인 콜옵션 순매수 수량’, ‘개인 풋옵션 순매수 수량’ 6개의 Data는 09:00분 ~ 09:30분까지의 30분간의 수치를 나타내고, ‘미결제약정’은 전일 폐장 시의 청산되지 않은 선물 계약 수이다. 마지막으로 ‘VIX 전일증가 대비 당일 증가’는 한국시각으로 장이 열리기 전 새벽에 마감되는 미국의 VIX 지수의 전일 증가대비 당일 증가의 변동치이다. 한국의 VKOSPI의 일 중 움직임은 VIX에 상당한 영향을 받기 때문에 VIX Data는 의미가 큰 Input 변수이다.

데이터는 NH투자증권에서 제공하는 NH Trader에서 받아왔으며, 2013년 11월 27일부터 2016년 5월 4일까지 거래일 중에서 일 중 VKOSPI가 상승한 날과 하락한 날의 비율을 정확하게 1:1로 맞추어 428개 씩(상승한 날 : 214일, 하락한 날 : 214일) 6회 무작위로 추출(Random Sampling)하였다. 이후 328개의 데이터 예제는 모델링을 위한 훈련데이터(Training Data)로 사용하였고, 나머지 100개의 데이터는 검증데이터(Test Data)로 사용하였다.

이때, 훈련데이터와 검증데이터 내의 상승한 날과 하락한 날의 비율을 1:1로 맞추었는데 그 이유는 공정한 모델링을 위해서이다. 만약 훈련데이터에서 VKOSPI가 상승한 경우의 데이터가 많다면 모델 자체가 VKOSPI 상승을 더 높은 확률로 예측하도록 치우칠 염려가 있기 때문이다. 데이터를 받아온 기간은 위에 명시했지만, 굳이 시간 순서를 맞춰줄 필요는 없으므로 시간순서와 상관없이 데이터를 추출하였다. 훈련데이터는 SVM 모델링을 하는 데에 사용하였고, 검증데이터는 예측값과 실제 값을 비교하는 임무를 수행하였다. <Table 3>에는 데이터 예제의 개수와 훈련데이터, 검증데이터로 나눈 개수를 쉽게 나타내었다.

<Table 3> Data partitioning and number of data

	Role	Number of days	
Data	Training data	328 (Up : 164, Down : 164)	428 (Up : 214, Down : 214)
	Test data	100 (Up : 50, Down : 50)	

3.2 SVM 모델링 결과

<Table 3>에 나타난 조건에 알맞게 6회 무작위로 추출하고 6회에 걸쳐 각 훈련데이터를 기반으로 모델링 하였다. 커널 함수는 ‘radial basis’를 사용했다. SVM 모델을 바탕으로 100개의 검증데이터의 Output 등락을 예측하고 실제 Output과 비교했다. 실제 일 중 VKOSPI 상승을 미리 상승으로 예측한 일 수와 마찬가지로 실제 하락을 하락으로 예측한 일 수를 더하여 그 합을 100으로 나누고, 이를 예측률이라고 정의하였다. <Table 4>에는 각 모델링의 SVM 설정 매개변수(Parameter)와 결과 테이블, 그리고 예측률이 나타나 있다. 결과 테이블의 가로축은 실제의 등락을, 세로축은 모델의 예측 등락을 나타내었다. 예측률은 결과 테이블의 대각 성분의 합으로서 예측률 평균값은 57.83%이었다.

<Table 4> SVM modeling results

Random sampling order	Parameters		Result table			Accuracy
	Cost	Gamma		U	D	
1	1	0.125		U	D	57%
			U	32	25	
			D	18	25	
2	1	1		U	D	59%
			U	24	15	
			D	26	35	
3	1	0.125		U	D	59%
			U	36	27	
			D	14	23	
4	1	1		U	D	58%
			U	26	18	
			D	24	32	
5	0.5	1		U	D	57%
			U	19	12	
			D	31	38	
6	0.5	1		U	D	57%
			U	16	9	
			D	34	41	

4. 일 중 VKOSPI 예측을 활용한 실제 옵션 매매에 적용

옵션이란 사전에 합의한 조건에 따라 약정기간에 해당 자산을 사거나 혹은 팔 수 있는 권리를 뜻한다. 특정 자산을 살 수 있는 권리를 가지는 것을 콜옵션, 팔 수 있는 권리를 가지는 것을 풋옵션이라고 한다. 옵션 매매는 단순한 주식거래나 선물거래보다 시장 상황에 따라 더욱 다양한 전략을 구사할 수 있는 특징을 가지고 있다.

4.1 옵션 양 매도 데이트레이딩 전략

‘양 매도’란 콜옵션과 풋옵션을 동시에 매도한다는 의미이다. 그리고 ‘데이트레이딩’이란 ‘당일 매매’라고 하며 상품을 사거나 파는 등 포지션을 가져간 이후 당일 안에 청산함으로써 장 마감 때까지 어떠한 포지션도 가져가지 않는 것을 의미한다. 즉 ‘옵션 양 매도 데이트레이딩 전략’이란 장 중의 어느 시점에 콜옵션과 풋옵션을 동시에 매도한 이후 그 날 안에 모든 포지션을 청산하는 전략이다. 일 중 VKOSPI의 움직임을 알고 있다면 이 전략에서 큰 수익을 낼 수 있다. (Bae, 2011)

많은 옵션투자자는 옵션 매매를 할 때 <Table 5>에 명시된 그릭스(Greeks)를 이용한다. 블랙-

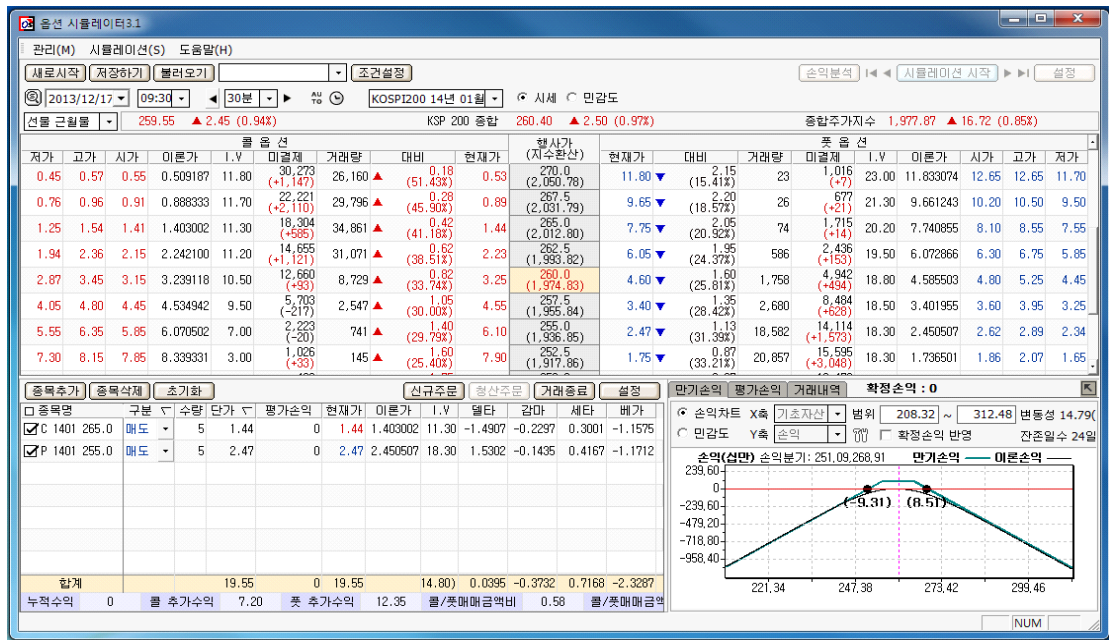
숄츠의 옵션 가격 결정 모형에서는 옵션의 가격을 결정하는 요소로 S(기초자산가격), K(행사가격), r(이자율), T(남은 기간), σ (변동성)를 들고 있다. 그릭스는 옵션 가격을 이 요소들로 미분한 것들이다.

그릭스 중 베가에서 언급한 변동성 σ 를 VKOSPI라고 이해해도 무방하다. 이 전략은 베가를 이용하여 수익을 낸다. 모든 옵션은 행사될 가능성이 커지는 상황일수록 옵션 가격이 오른다. 변동성이 오르면 기초자산의 기대 가격변화 폭이 커지고 이는 만기시점에서의 옵션의 행사 가능성을 더 높인다. 따라서 옵션 시장에서는 변동성이 증가할 때에는 옵션 가격이 오르고, 실제로 콜옵션, 풋옵션 구분 없이 모든 옵션의 베가는 항상 (+)값을 가진다. 즉 변동성 상승이 옵션 가격의 상승을 유발한다는 것이다. 옵션을 매수한 상태에서 변동성이 오른다면 베가에 의해 옵션 가격도 오르게 되고 이는 포지션에 수익이 나게 된다. 반대로 옵션을 매도한 상태였다면 옵션 가격이 오르는 과정에서 포지션은 손실을 보게 된다. 반면 변동성이 떨어지는 상황에서는 베가에 의한 수익이 나게 된다.

양 매도 구조상 일 중 VKOSPI가 하락한다면 베가의 측면에서는 수익이 나게 된다. 물론 포지션의 수익은 베가에 의해서만 결정되는 것이 아니고 나머지 4개의 그릭스도 고려해줘야 한다.

<Table 5> Greeks of options

Delta	Change in option prices for change in underlying asset prices.
Gamma	Change in delta for change in underlying asset prices. The second derivative of the option price for the price change of the underlying asset.
Theta	Change in option prices for change in time.
Vega	Change in option prices for change in volatility(σ).
Rho	Change in option prices for change in risk free rate.



〈Figure 4〉 Strangle sell in option simulator 3.1

만약 배가를 제외한 나머지 그리스가 적절히 통제된다면 배가가 포지션 수익변화에 가장 큰 비중을 차지할 것이고 결국은 배가에서만 수익이 난다고 하여도 무방하다.

나머지 그리스를 통제하는 방법으로는 처음으로 델타와 감마는 포지션 델타값을 0에 가깝게 맞추는 델타 중립(Delta Neutral)을 만들며 통제한다. 각 옵션은 종류별로 델타값이 다르다. 콜 옵션의 델타는 0~1에 위치하고, 풋옵션의 델타는 -1~0에 위치한다. 옵션의 행사가격과 수량조절을 적절히 한다면 델타를 거의 0에 근접하도록 양 매도 포지션을 구축할 수 있다. 델타 중립이 이루어진다면 장 중에 적당한 기초자산 가격의 변화에 대해서는 옵션 가격이 크게 변하지 않는다. 이로써 델타와 감마를 통제할 수 있다. 다만 기초자산의 가격이 크게 움직이는 날에는 감

마 때문에 계속 델타값이 한쪽으로 치우칠 수 있으므로 주의해야 한다.

세타와 로우는 특별히 투자자가 통제를 할 필요는 없다. 세타의 경우에는 데이트레이딩이기 때문에 하루 만에 세타에 의해 옵션 가격이 크게 변하지 않는다. 세타는 주로 오랜 기간 포지션을 가져가는 포지션 매매 시 고려하는 부분이다. 데이트레이딩이라는 단기간 매매방식 자체가 세타를 자동으로 통제해준다고 보면 된다. 또한, 로우의 경우에도 무위험자산의 이자율은 하루아침에 크게 변하는 값이 아니므로 별다른 통제를 할 필요는 없다.

실제로 일 중 VKOSPI 예측이 100% 가능하다는 가정 아래에 일 중 VKOSPI가 하락하는 날의 옵션 양 매도 데이트레이딩 전략의 성과가 있을 것인가에 대한 실험을 해보았다. 2013년 11월 27

일부터 2016년 5월 4일까지 거래일 중 미국 시장과 한국 시장의 부득이한 이유로 인하여 VIX 데이터를 Input으로 얻을 수 없는 날을 제외하고 총 556거래일에 해당하는 기간을 대상으로 하였다. 옵션 양 매도 포지션을 09:30분에 포지션을 가져가며, 정확히 15:00분에 포지션 청산을 한다. <Figure 4>는 시뮬레이션에 사용된 프로그램으로서 옵션 시뮬레이터 3.1버전³⁾이다.

포지션 구축 시 거래한 모든 옵션은 행사가격이 ATM(등가격 옵션) ~ OTM(외가격 옵션) 3단계 사이에 해당하는 것들이다. 일반적으로 ITM(내가격 옵션)보다 OTM 옵션이 가격이 낮아 거래량이 많으므로 원활한 거래 체결을 위해서 행사가격의 제한을 두며 포지션을 구축하였다. 또한, 계약 수는 수량 조절을 통한 델타 중립을 위해 콜옵션, 풋옵션 각각 5개 내외로 하였으며 포

지션 델타의 절댓값이 0.05를 넘지 않도록 포지션 구축을 하였다. <Table 6>은 556일에 대한 시뮬레이션 결과의 일부이다.

<Table 7>은 일 중 VKOSPI의 실제 등락을 기준으로 구분된 표이다. 일 중 VKOSPI가 상승한 날에는 옵션 양 매도 데이트레이딩 전략 실행이 손실로 이어지는 경우가 많았고, 반대로 일 중 VKOSPI가 하락한 날에는 전략 실행이 수익으로 이어지는 경우가 손실로 이어지는 경우보다 4배 가까이 많았다. 이로써 일 중 VKOSPI의 하락이 전략 실행에서 수익을 가져다준다는 것을 알 수 있으며, 전략의 타당성이 검증되었다.

<Figure 5> 그래프의 우상향하는 실선은 실제 일 중 VKOSPI가 하락한 날에만 전략을 실행했다고 가정한 경우이고, 점선은 일 중 VKOSPI 등락과 관계없이 전략을 실행했다고 가정한 경우

<Table 6> Part of the strangle sell simulation

Date	Output	Call	Number of contracts	Put	Number of contracts	Position delta	Return (\)
2015-08-04	Down	O-2	5	O-2	5	0.0269	275,000
2015-08-05	Down	O-2	6	O-1	3	0.0254	165,000
2015-08-06	Up	O-2	4	O-2	6	0.0125	-1,190,000
2015-08-07	Down	O-2	5	O-3	5	0.0383	175,000
2015-08-10	Down	O-2	5	O-3	5	-0.0216	-25,000
2015-08-11	Up	O-2	5	O-2	5	0.0495	-300,000
2015-08-12	Up	O-1	6	O-1	7	0.0272	-175,000
2015-08-13	Down	O-1	8	O-1	3	0.0088	430,000
2015-08-17	Down	O-1	4	O-1	5	-0.0264	125,000
2015-08-18	Down	O-2	3	O-2	4	0.0134	115,000
2015-08-19	Up	O-1	5	O-1	5	-0.0201	-375,000

* Ignore commissions and slippage costs.

* In the Output section, 'Up' is when VKOSPI rose during the day and 'Down' is when VKOSPI dropped during the day.

* O-n : Options of out the money.

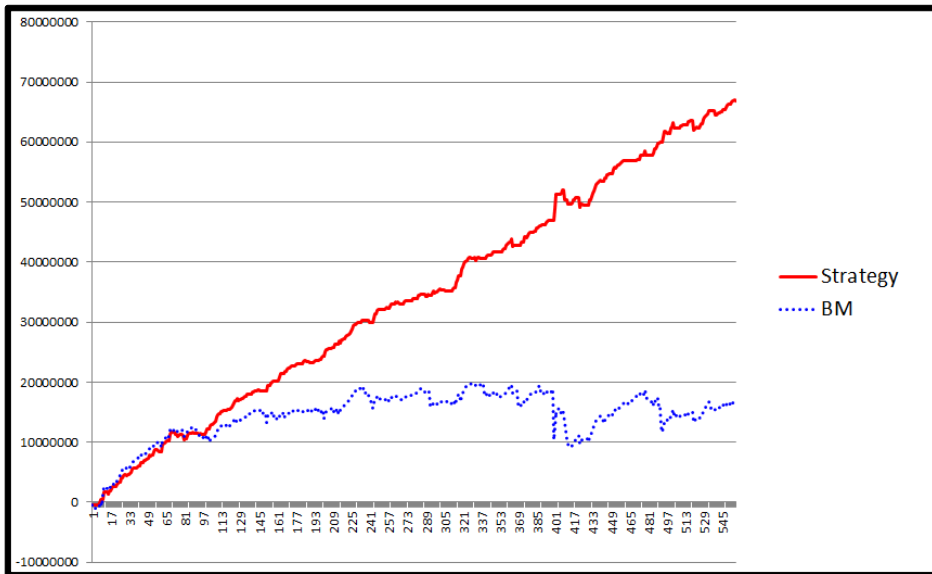
* Position delta : Delta of strangle sell position.

3) 예스스탁(www.yesstock.com)에서 제공하는 옵션 매매 시뮬레이션 프로그램.

〈Table 7〉 Performance of strategy based on actual VKOSPI change during the day

Total date	Actual VKOSPI change during the day		Performance of daily strangle sell strategy	
	556	Up	214일	Plus return
Minus return				122
Return = 0				2
Down		342일	Plus return	269
			Minus return	69
			Return = 0	4

* Ignore commissions and slippage costs.



* Ignore commissions and slippage costs.

〈Figure 5〉 Performance graph of strategy based on actual VKOSPI change during the day

이며 벤치마크 지표로 사용한다. 모델을 기반으로 전략을 실행한 경우 약 6,700만 원의 수익이 발생하고, 이는 약 1,700만 원의 수익이 발생한 벤치마크 지표와 대비해 4배에 달하는 수익을 가진다.

하지만 실제 일 중 VKOSPI의 변화를 09:30분에 미리 알 수 없으므로 <Figure 5>에 나타난 수

익 곡선을 실현할 수는 없지만 높은 확률의 일 중 VKOSPI 예측 능력이 있다면 충분한 수익을 올릴 수 있다.

4.2 SVM 모델 기반 전략 실험 결과

3.2절에서는 SVM 모델을 통해 일 중 VKOSPI

를 예측해 보았다. 이후 모델이 검증데이터의 Output을 하락으로 예측한 경우에 대해서만 옵션 양 매도 데이트레이딩 전략을 실행했다. 비교 대상 선정을 위해 전체 100거래일 모두 해당 전략을 실행한 경우의 결과를 벤치마크 지표로 두었다. 이후 SVM 모델 기반 전략의 성과와 벤치마크 지표의 성과를 비교해 보았고, <Table 8>은 6개의 모델과 모델별 벤치마크 지표 6개의 평균을 각각 손익, 거래 횟수로 비교한 결과를 보여준다.

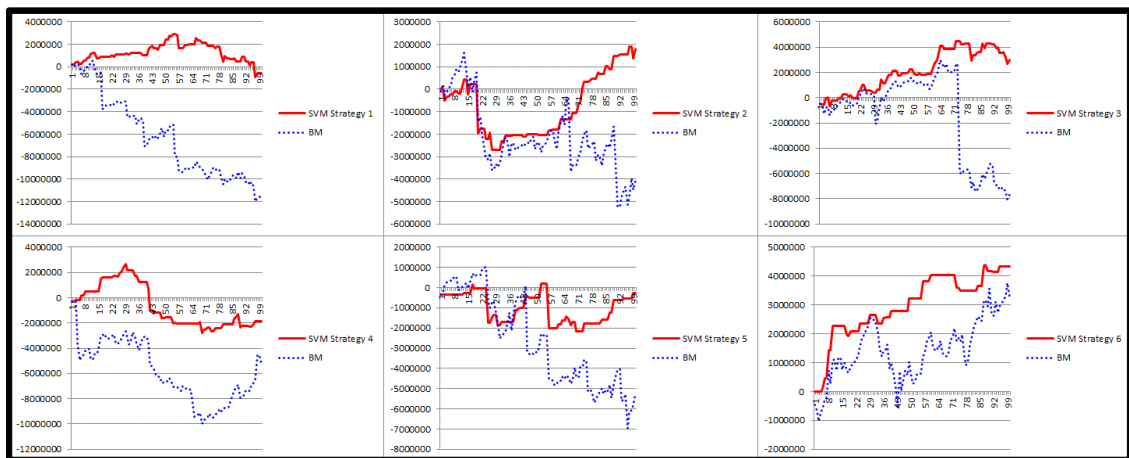
SVM 모델을 기반으로 전략을 실행한 6개의

경우 중에서 가장 큰 수익이 발생한 경우는 6번째 경우이며 4,340,000원의 수익이 발생하였다. 반면 벤치마크의 평균 손익은 -5,085,000원이며 이는 일 중 VKOSPI를 고려하지 않은 전략의 실행은 손실 확률이 높다는 것을 의미한다. 또한, 각각의 경우에 모두 각자의 벤치마크보다 수익이 높은 것을 <Figure 6>를 통해 알 수 있다. 거래 횟수도 평균 약 43.2회로써 벤치마크 평균인 100회보다 절반 이상 낮다. 이로써 거래 횟수 측면에서 더 효율적인 매매를 하였다고 판단된다.

<Table 8> Performance comparison between strategy based on each SVM model and benchmark average

	Return (₩)	Number of transactions
SVM Strategy 1	-600,000	57
SVM Strategy 2	1,790,000	39
SVM Strategy 3	3,010,000	63
SVM Strategy 4	-1,860,000	44
SVM Strategy 5	-260,000	31
SVM Strategy 6	4,340,000	25
BM average	-5,085,000	100

* Ignore commissions and slippage costs.



<Figure 6> Performance graph of strategy based on each SVM model

5. 결론 및 향후연구

변동성은 금융 시장에서 여러 상품을 매매하는 투자자에게 상당히 중요하다. 그동안의 많은 연구에서는 다양한 방법으로 변동성 예측을 시도하였다(Dimpfl and Jank, 2013; Ghysels and Sohn, 2009).

본 연구는 변동성 예측을 위해 기계학습 방법을 이용하였기에 희소가치가 높다. 변동성 예측 후 옵션 매매에 대한 적용 아이디어를 제시하면서 매매성과를 높일 아이디어와 긍정적인 시뮬레이션 결과를 제시하였다. 옵션은 주식, 선물보다 레버리지 효과가 훨씬 크기 때문에 옵션에 대한 지식 없이 매매하는 투자자들에게는 큰 손실을 안겨줄 수 있는 위험한 파생상품이다. 하지만 반대로 옵션의 구조와 시장 상황에 관해 높은 이해도가 바탕이 되어 있고, 이를 이용할 수 있는 투자자들에게는 투자가치가 높은 상품이다. 본 연구는 투자성 높은 상품에 대한 매매전략을 연구했기 때문에 또한 높은 가치가 있다고 판단한다.

본 연구의 한계점은 VKOSPI의 예측력이다. VKOSPI의 예측력이 정확할수록 옵션 매매 성과가 높아짐을 연구를 통해 밝혔다. 따라서 VKOSPI 예측력을 한층 더 높임으로써 매매 성과를 더 높일 수 있도록 보완할 필요가 있다. 더욱더 정교한 SVM 모델을 고안하는 것이 합리적인 방법이다. 추가 연구를 통해 예측력이 높아진다면 더 좋은 옵션 매매 성과를 기대해 볼 수 있을 것이다.

참고문헌(References)

- Bae, M. K., "Profitability of Intra-day Short Volatility Strategy using Volatility Risk Premium," *Master's dissertations*, The Graduate School of Business IT, Kookmin University, 2011.
- Boser, B. E. and I. M. Guyon and V. N. Vapnik, *A Training Algorithm for Optimal Margin Classifiers*, COLT '92 Proceedings of the fifth annual workshop on Computational learning theory, 1992. Available at <http://dl.acm.org/citation.cfm?doid=130385.130401> (Accessed 8 September, 2016).
- Cho, D., "Implied Volatilities of the KOSPI 200 Index Option Market," *Korean Journal of Futures and Options*, Vol.23, No.4(2015), 517~541.
- Choi, Y. S. and H. J. Lee, "Forecasting KOSPI 200 Volatility by Volatility Measurements," *Communications of the Korean Statistical Society*, Vol.17, No.2(2010), 293~308.
- Choi, H. S. and S. W. Kim and S. C. Park, "Analysis of Trading Performance on Intelligent Trading System for Directional Trading," *Journal of Intelligence and Information Systems*, Vol.17, No.3(2011), 187~201.
- Dimpfl, T. and S., Jank, "Can Internet Search Queries Help to Predict Stock Market Volatility?," *European Financial Management*, Vol.22, No.2(2016), 171~192.
- Flach, P., *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*, Cambridge University Press, 2012.
- Ghysels, E. and B., Sohn, "Which power variation predicts volatility well?," *Journal of*

- Empirical Finance*. Vol.16, No.4(2009), 686~700.
- Jeong, J. H. and J. O. Lim and S. Y. Jei, "Comparative Study of the Spill-over in the Conditional Volatility and Dynamic Conditional Correlation of Various Stock Markets," *The Korean Journal of Financial Engineering*, Vol.11, No.1(2012), 1~16.
- Jeong, Y. G. and Y. S. Yun, "A Study on the Predictability of Stock Price Using Artificial Neural Network Model," *The Korean Journal of Finance Management*, Vol.15, No.2(1998), 369~399.
- Kim, S. W., "A Study on Developing a VKOSPI Forecasting Model via GARCH Class Models for Intelligent Volatility Trading Systems," *Journal of Intelligence and Information Systems*, Vol.16, No.2(2010), 19~32.
- Kim, S. W. and H. C. Ahn, "Development of an Intelligent Trading System using Support Vector Machines and Genetic Algorithms," *Journal of Intelligence and Information Systems*, Vol.16, No.1(2010), 71~92.
- Lantz, B., *Machine Learning with R*. Packt, Birmingham, UK, 2013.
- Li, X. and D. Lord and Y. Zhang and Y. Xie, "Predicting motor vehicle crashes using Support Vector Machine models," *Accident Analysis & Prevention*, Vol.40, No.4(2008), 1611~1618.
- Lim, H. C. and Y. S. Choi, "Knock-In and Stocks Market Effect Due to ELS Issuance and Hedging," *Korean Journal of Futures and Options*, Vol.23, No.2(2015), 289~321.
- Moon, J. B., "A Study on VKOSPI Volatility Trading Strategy using VIX Index," *Master's dissertations*, The Graduate School of Business IT, Kookmin University, 2016.
- Oh, I. S., *Pattern Recognition*, Kyobobook, 2008.
- Park, S. C. and S. W. Kim and H. S. Choi, "Selection Model of System Trading Strategies using SVM," *Journal of Intelligence and Information Systems*, Vol.20, No.2(2014), 59~71.
- Park, S. H. and B. Hansen, "Prediction of Protein-Protein Interaction Sites Based on 3D Surface Patches Using SVM," *Korea Information Processing Society*, Vol.19, No.1(2012), 21~28.
- Rajhans, R. K. and A. Jain, "Volatility Spillover," *Paradigm*, Vol.19, No.2(2015), 137~151.
- Sharma, R. K., "Developing Models to Predict the Stock Prices of Indian Automobile Giant Tata Motors Company Limited: An Empirical Analysis," *Finance India*, Vol.30, No.1(2016), 153~172.
- Siddiqui, T. A. and Y. Abdullah, "Developing a Nonlinear Model to Predict Stock Prices in India: An Artificial Neural Network Approach," *IUP Journal of Applied Finance*, Vol.21, No.1(2015), 36~49.
- Tay, F. E. H. and L. J. Cao, "Modified support vector machines in financial time series forecasting," *Neurocomputing*, Vol.48, No.1-4(2002), 847~861.

Abstract

VKOSPI Forecasting and Option Trading Application Using SVM

Yun Seon Ra* · Heung Sik Choi** · Sun Woong Kim***

Machine learning is a field of artificial intelligence. It refers to an area of computer science related to providing machines the ability to perform their own data analysis, decision making and forecasting. For example, one of the representative machine learning models is artificial neural network, which is a statistical learning algorithm inspired by the neural network structure of biology. In addition, there are other machine learning models such as decision tree model, naive bayes model and SVM(support vector machine) model. Among the machine learning models, we use SVM model in this study because it is mainly used for classification and regression analysis that fits well to our study. The core principle of SVM is to find a reasonable hyperplane that distinguishes different group in the data space. Given information about the data in any two groups, the SVM model judges to which group the new data belongs based on the hyperplane obtained from the given data set. Thus, the more the amount of meaningful data, the better the machine learning ability.

In recent years, many financial experts have focused on machine learning, seeing the possibility of combining with machine learning and the financial field where vast amounts of financial data exist. Machine learning techniques have been proved to be powerful in describing the non-stationary and chaotic stock price dynamics. A lot of researches have been successfully conducted on forecasting of stock prices using machine learning algorithms. Recently, financial companies have begun to provide Robo-Advisor service, a compound word of Robot and Advisor, which can perform various financial tasks through advanced algorithms using rapidly changing huge amount of data. Robo-Advisor's main task is to advise the investors about the investor's personal investment propensity and to provide the service to manage the portfolio automatically.

* Graduate School of Business IT, Kookmin University

** Corresponding Author: Heung Sik Choi

Graduate School of Business IT, Kookmin University

77 Jeongneung-ro, Seongbuk-gu, Seoul 02707, Korea

Tel: +82-2-910-4567, Fax: +82-2-910-4017 E-mail: hschoi@kookmin.ac.kr

*** Graduate School of Business IT, Kookmin University

In this study, we propose a method of forecasting the Korean volatility index, VKOSPI, using the SVM model, which is one of the machine learning methods, and applying it to real option trading to increase the trading performance. VKOSPI is a measure of the future volatility of the KOSPI 200 index based on KOSPI 200 index option prices. VKOSPI is similar to the VIX index, which is based on S&P 500 option price in the United States. The Korea Exchange(KRX) calculates and announce the real-time VKOSPI index. VKOSPI is the same as the usual volatility and affects the option prices. The direction of VKOSPI and option prices show positive relation regardless of the option type (call and put options with various striking prices). If the volatility increases, all of the call and put option premium increases because the probability of the option's exercise possibility increases. The investor can know the rising value of the option price with respect to the volatility rising value in real time through Vega, a Black-Scholes's measurement index of an option's sensitivity to changes in the volatility. Therefore, accurate forecasting of VKOSPI movements is one of the important factors that can generate profit in option trading.

In this study, we verified through real option data that the accurate forecast of VKOSPI is able to make a big profit in real option trading. To the best of our knowledge, there have been no studies on the idea of predicting the direction of VKOSPI based on machine learning and introducing the idea of applying it to actual option trading.

In this study predicted daily VKOSPI changes through SVM model and then made intraday option strangle position, which gives profit as option prices reduce, only when VKOSPI is expected to decline during daytime. We analyzed the results and tested whether it is applicable to real option trading based on SVM's prediction. The results showed the prediction accuracy of VKOSPI was 57.83% on average, and the number of position entry times was 43.2 times, which is less than half of the benchmark (100 times). A small number of trading is an indicator of trading efficiency. In addition, the experiment proved that the trading performance was significantly higher than the benchmark.

Key Words : Machine Learning, Support Vector Machine, VKOSPI, Option Trading

Received : November 1, 2016 Revised : December 13, 2016 Accepted : December 16, 2016

Publication Type : Regular Paper Corresponding Author : Heung Sik Choi

저 자 소개



리운선

국민대학교 수학과에서 이학학사를 취득하였고, 현재 국민대학교 비즈니스IT전문대학원에서 트레이딩시스템 전공 석사 과정에 재학 중이다. 관심분야로 시스템 트레이딩, 머신 러닝, 금융공학 등이 있다.



최흥식

현재 국민대학교 경영대학 경영정보학부 및 동 대학 비즈니스IT전문대학원 교수로 재직 중이다. KAIST에서 경영과학 석사학위를 취득하였으며 미국 로체스터 대학에서 경영학 석사 및 박사학위를 취득하였다. 관심분야로는 파생상품 시스템트레이딩, 트레이딩계량 분석, 옵션 변동성매매 등이다.



김선웅

현재 국민대학교 비즈니스IT전문대학원 교수로 재직 중이다. 서울대학교 경영학과에서 경영학사를 취득하고, KAIST 경영과학과에서 투자론을 전공하여 공학석사와 공학 박사학위를 취득하였다. 주요 관심분야는 트레이딩시스템, 투자공학, 헤지펀드와 자산 운용이다.