

# 주가 경향 예측 모델의 공정한 성능 평가 방법

## Fair Performance Evaluation Method for Stock Trend Prediction Models

임정수

한국교통대학교 전자공학과

Chungsoo Lim(clim@ut.ac.kr)

### 요약

주식 투자는 재테크의 하나로 금리 인하와 비과세 제도의 축소에 따라 주목을 받기 시작했다. 그러나 투자에 전문적인 지식이 필요할 뿐 아니라 위험 부담이 크다는 단점이 있다. 따라서 주가 경향의 정확한 예측은 개인 투자자에게나 주식 투자 관련 서비스를 제공하는 회사에 중요한 능력이며, 더욱 정확한 예측을 위한 연구가 활발히 진행 중이다. 그러나 예측 연구들의 공정한 비교와 최고의 예측 모델을 얻기 위한 하이퍼-파라미터의 최적화에는 예측 모델의 성능을 정확하게 평가하는 방법이 필요한데, 지금까지 예측 모델의 성능 평가에 대한 연구는 미진한 상태이며, 기존 방법들을 그대로 답습하고 있는 실정이다. 이에 본 논문에서는 주가 예측 모델 성능 평가를 측정기준과 데이터 구성의 관점에서 분석하고, 예측 불균형 비율을 이용한 주가 경향 예측 모델의 공정한 성능 평가 방법을 제안한다.

■ 중심어 : | 주가 경향 예측 | 주가 경향 예측 성능 평가 방법 | 기계학습 |

### Abstract

Stock investment is a personal investment technique that has gathered tremendous interest since the reduction in interest rates and tax exemption. However, it is risky especially for those who do not have expert knowledge on stock volatility. Therefore, it is well understood that accurate stock trend prediction can greatly help stock investment, giving birth to a volume of research work in the field.

In order to compare different research works and to optimize hyper-parameters for prediction models, it is required to have an evaluation standard that can accurately assess performances of prediction models. However, little research has been done in the area, and conventionally used methods have been employed repeatedly without being rigorously validated. For this reason, we first analyze performance evaluation of stock trend prediction with respect to performance metrics and data composition, and propose a fair evaluation method based on prediction disparity ratio.

■ keyword : | Stock Price Trend Prediction | Performance Evaluation Method for Stock Trend Prediction | Machine Learning |

## I. 서론

경제성장으로 인한 높아진 생활수준, 세계화

(globalization), 정보통신기술 (information and communication technology)의 발전, 금리 인하와 비과세 제도의 축소 등으로 인해 사람들은 간편한 투자

접수일자 : 2020년 07월 22일

수정일자 : 2020년 08월 25일

심사완료일 : 2020년 08월 25일

교신저자 : 임정수, e-mail : clim@ut.ac.kr

로 높은 수익을 얻을 수 있는 주식에 점차 관심을 가지게 되었다[1]. 또한 국내의 주식시장도 성장하여 1996년 상장회사 수 760개, 시가총액 117.4조원이던 것이 2019년 상장회사 수 2204개, 시가총액 1717조원이 되었다[2]. 그러나 주식 투자는 높은 위험성 (high risk)을 동반하며, 만약 주가의 변동에 대한 이해부족, 심적 회계 (mental accounting), 지나친 자신감 (overconfidence) 등에 기인한 비합리적인 투자가 이루어진다면 자신이 축적한 부를 모두 잃을 수도 있다 [3]. 그러므로 주식 투자자들의 합리적인 판단을 도와 최대의 수익을 얻게 하는 거래시스템 (trading system)의 필요성이 높아지게 되었다. 경제와 투자 관련 알고리즘과 빅데이터 분석에 기반한 온라인 재무 및 투자 관련 서비스를 로보어드바이저 (robo-advisor)라고 하는데, 2017년 운용 자산 규모로 가장 큰 로보어드바이저 회사는 The Vanguard Group으로 830억 달러를 운용하였고[4], 우리나라의 로보어드바이저 회사인 신한에이아이는 2020년 운용 자산 규모가 1500억 원 정도이다[5]. 아직 운용 자산 규모가 크지 않지만, 인공지능의 발전으로 보다 높은 수익이 보장된다면 그 규모가 점차 커질 것으로 예상된다. 이런 거래시스템의 성능은 국내의 모든 수익률을 가지고 나타내는데, 이 수익률은 주가의 미래 값이나 경향 예측의 정확도에 크게 영향을 받기 때문에 주가의 예측에 대한 연구가 많이 진행되고 있다[6].

주가와 관련된 예측에는 예측 대상에 따라 두 가지 종류로 나눌 수가 있다. 미래의 주가 자체의 예측과 미래 주가의 경향 (trend) 예측이 그것이다. 이 예측들에 근거하여 거래 신호 (매입, 보유, 매도) 발생, 거래 시점 예측 등이 가능하다. 지금까지 주가 자체를 예측하는 연구가 보다 많이 수행되었지만, 주가 경향을 예측하는 것이 주가 자체를 예측하는 것보다 주가 수익에 더욱 도움이 되며 더욱 정확한 거래 결과를 얻을 수 있다는 연구도 발표되었다[7].

주가 관련 예측의 정확도를 높이고 주식거래의 질을 높여 보다 높은 수익을 얻기 위해 주로 두 가지 종류의 분석방법이 사용된다. 첫 번째 방법은 기본적 분석 (fundamental analysis)으로 주식의 내재적 가치를 분석하여 주가를 예측하는 방법으로, 해당 회사의 재무

제표, 산업계과 경제의 상황, 거시경제 환경 등을 고려한다. 두 번째 방법은 기술적 분석 (technical analysis)으로 주가에 영향을 주는 요인들이 주가, 거래량 등의 값에 집약되어 있다고 보고 이들을 분석하여 패턴을 찾아내고 이를 이용하여 주가를 예측하는 방법이다. 분석을 위해 주가와 거래량으로부터 여러 가지 기술지표를 구하여 사용한다.

이러한 기본적 분석과 기술적 분석법들은 계속 주가 예측에 사용되고 있다. 그 이유는 효율적 시장 가설 (efficient-market hypothesis)이 완벽하게 적용되지 않는 시장이 존재하며, 랜덤 또는 비선형 데이터에서 패턴을 찾아낼 수 있는 기계학습 (machine learning) 기법들이 컴퓨터 기술의 발전으로 그 활용 범위가 넓어졌기 때문이다[7][8]. 실제로 기본적 분석, 기술적 분석은 기계학습 기법과 함께 전문 트레이더들에 의해 사용되고 있으며, 기술분석과 기계학습 조합이 자주 사용되고 있다[9].

이렇게 기계학습을 이용한 많은 주가 관련 예측 연구가 발표되고 있으므로 각 연구에서 제안하는 예측 모델들의 비교를 위해서는 각 모델의 예측 성능을 평가하는 공통된 방법이 필요하다. 또한, 기계학습을 이용하여 예측 모델을 만들 때 거쳐야 하는 하이퍼-파라미터 (hyper-parameter) 최적화 과정에서도 모델의 예측 성능을 평가하는 방법이 필요하다. 그러나 지금까지 대부분의 연구에서는 기존에 사용되고 있던 평가 방법을 그 타당성에 대한 아무 검증 없이 사용해왔다. 이 평가 방법의 문제점은 주로 정확도만을 측정기준 (metric)으로 사용하며, 테스트 데이터 (test data)와 검증 데이터 (validation data)의 레이블 ('상승'과 '하강') 구성에 관한 기준이 없다는 것이다. 이 문제들은 이진 분류 (binary classification)에 있어서 공통적인 문제들이지만 특히 주식의 경향 예측에 더 큰 문제가 될 수 있다. 그 이유는 주식 경향 예측의 정확도는 대부분 50%에서 70% 사이로 상대적으로 낮은 편이라 데이터의 구성에 따라 예측 모델의 성능이 과장될 수 있기 때문이다. 예를 들어 클래스가 '상승'인 벡터가 전체 테스트 데이터의 60%를 차지하고, '하강'인 벡터가 40%를 차지한다고 했을 때 모든 테스트 데이터 벡터를 '상승'으로 예측하는 모델의 정확도는 60%가 되고, 심각한 예측의

불균형에도 불구하고 정확도만 고려된다면 이 모델이 가장 좋은 모델로 선택될 수도 있다.

따라서 본 논문에서는 첫 번째로 기존에 사용해오던 주가 경향 예측 모델의 성능 평가 방법에 문제가 없는지 성능 측정기준과 데이터 구성의 관점에서 분석한다. 그리고 두 번째로 이 분석 결과를 바탕으로 새로운 성능 평가 방법을 제안한다.

## II. 연구 배경

이번 장에서는 본 논문에서 수행하고자 하는 두 가지 분석에 대한 배경을 설명한다. [표 1]은 주식 경향의 예측을 주제로 하는 논문에서 사용한 측정기준 (metrics) 과 데이터 구성 (data composition)을 정리해 놓은 것이다.

먼저 측정기준을 살펴보면 다양한 측정기준이 존재

하는 것을 알 수 있다. 정확도 (accuracy), 재현율 (recall), 특이도 (specificity), 정밀도 (precision), 음성예측도 (negative predictive value, NPV), F-score 등의 예측 성능을 나타내는 지표뿐 아니라 수익과 관련된 수익률 (return rate)과 샤프 비율 (Sharpe ratio) 등이 사용되고 있다. 여기서 재현율, 특이도, 정밀도, 음성예측도는 '상승'을 양성 (positive)로, '하강'을 음성 (negative)로 정의할 때의 측정기준이다. 이 중 정확도는 거의 모든 연구에서 사용되며 정확도만 사용하는 연구도 있다. 하지만 최근의 연구에서는 정확도뿐 아니라 다른 측정기준도 함께 사용하는 경향이 있다. 그러나 두 개 또는 세 개의 측정기준을 사용하는 경우가 대부분이며, 재현율, 특이도, 정밀도, 음성예측도를 모두 제시하는 연구는 보이지 않는다. 이렇게 각 연구에서 사용하는 측정기준이 다르기 때문에 본 논문에서는 주가 경향 예측 모델의 성능을 온전히 나타내기 위해서 어떤 측정기준들이 필요한지 분석해 보고자 한다.

표 1. 주가 경향 예측 연구 비교 분석

Ref.	Year	Metrics	Data composition information
[1]	2019	hit ratio, recall, specificity	not presented
[3]	2014	accuracy, precision	presented (entire data)
[6]	2019	F-score, sharpe ratio	not presented
[7]	2016	accuracy, return rate	not presented
[8]	2018	accuracy, return per trade	not presented
[9]	2018	accuracy, profit rate	not presented
[10]	2010	accuracy	presented (entire data)
[11]	2013	accuracy	presented (test)
[12]	2015	accuracy, F-score	presented (training, test)
[13]	2018	accuracy	presented (training, validation, test)
[14]	2019	accuracy, profit rate	presented (training, test)
[15]	2010	area under ROC	not presented
[16]	2010	accuracy	not presented
[17]	2011	accuracy, recall, specificity	not presented
[18]	2011	accuracy	not presented
[19]	2015	accuracy	not presented
[20]	2015	accuracy	not presented
[21]	2013	accuracy	not presented
[22]	2014	accuracy	not presented
[23]	2013	accuracy	not presented
[24]	2003	accuracy	not presented
[25]	2013	accuracy	not presented
[26]	2014	accuracy, recall, specificity	not presented
[27]	2012	accuracy, profit rate	not presented
[28]	2014	accuracy	not presented
[29]	2000	accuracy	not presented
[30]	2016	accuracy, recall, specificity, sharpe ratio	not presented
[31]	2009	accuracy	not presented
[32]	2011	accuracy	not presented
[33]	2016	accuracy, recall, precision, F-score, joint prediction error	not presented

두 번째로 표에서 연구에 사용된 데이터, 즉 '상승'과 '하강'으로 레이블된 벡터들의 구성비율을 살펴보면, 학습 데이터, 검증 데이터, 테스트 데이터의 각 구성에 대해서는 대부분의 연구에서 밝히지 않고 있음을 알 수 있다. 모든 연구에서 전체 데이터의 크기와 학습 데이터, 검증 데이터 테스트 데이터의 각 크기를 언급하고 있는 것과 대조적이다. 데이터 구성에 대해 제시한 연구 중에는 전체 데이터의 '상승'과 '하강' 데이터 비율만 제시하여 학습, 검증, 테스트 데이터의 구성비율은 알 수 없는 경우도 있다[3][10]. 또한 학습, 검증, 테스트 데이터의 구성비율을 표시한 연구에서도 구성이 '상승'과 '하강'의 비율을 동일하게 맞추는 시도는 보이지 않는다[11-14]. 즉 모든 연구에서 데이터를 시간순으로 또는 랜덤하게 나누어 학습, 검증, 테스트 데이터를 구성하고 '상승'과 '하강'으로 레이블된 데이터의 비율은 크게 고려하지 않았기 때문에 이런 결과가 나온 것을 보인다. 따라서 본 논문에서는 데이터 구성 시에 '상승'과 '하강'으로 레이블된 데이터의 비율을 고려하지 않는 경우 어떤 영향이 있는지 분석하려고 한다.

### III. 분석

#### 1. 분석 환경

분석에 사용되는 주식 데이터에 대한 정보, 학습과 검증 데이터 분할, 특징벡터, 특징 선별법, 전처리, 그리고 주가 경향 예측에 사용된 분류기는 [표 2]에 정리되어 있다.

분석을 위해 대표적 주식 200개 종목으로 산출되는 주가지수인 KOSPI200을 비롯하여 삼성전자, 한국전력, 모비스, 기업은행의 10년치 주가를 이용하였다. 전체 데이터의 70%는 학습 데이터로, 그리고 나머지 30%는 검증 데이터로 사용하였으며 검증 데이터 내의 '상승'으로 레이블된 벡터와 '하강'으로 레이블된 벡터의 비율은 동일하게 설정하였다. 그리고 27개의 기술지표(technical indicator) 중 주성분 분석(principal component analysis)을 통하여 선별된 지표를 특징벡터로 사용하였다. 특징별로 정규화를 수행했으며 예측을 위한 분류기로는 서포트벡터머신(support

vector machine)을 사용하였다.

#### 2. 주가 경향 예측 성능의 측정기준 분석

주식 경향 예측에서 사용되는 측정기준은 제안되는 예측 방법들의 공정한 비교와 SVM, neural network 과 같은 기계학습 기법의 하이퍼-파라미터의 선택을 위해 예측 모델의 성능을 온전히 나타내야 한다. 그러나 [표 1]을 보면 측정기준으로 정확도만을 사용하는 연구도 많이 있으며 각 연구마다 사용하는 측정기준이 다르다. 따라서 본 절에서는 정확도, 재현율, 특이도, 정밀도, 음성예측도 중에서 예측의 성능을 온전히 나타내기 위해서 어떤 기준들의 사용되어야 하는지 분석하고자 한다.

표 2. 분석 환경

사용된 주식 (2009 ~ 2018)	KOSPI200, 삼성전자, 한국전력, 모비스, 기업은행
학습/분류데이터 분할	학습 데이터: 70%, 검증 데이터: 30% 검증 데이터 중 '상승', '하강'의 비율 동일하게 조정
특징 (27개)	시가, 종가, 상한가, 하한가, 거래량, stochastic %K, stochastic %D, slow stochastic %D, momentum, ROC, williams' %R, A/D oscillator, disparity5, disparity10, OSCP, CCI, RSI, ROV, EMA, MACD, MFI, MAO, VROC, MAs of [시가, 종가, 상한가, 하한가]
특징 선별법	주성분 분석 (principal component analysis)
전처리	각 특징별로 min-max normalization (-1 ~ 1)
분류기	SVM (kernel = RBF)

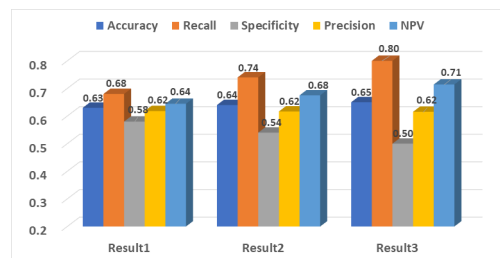


그림 1. 주가 경향 예측 결과 예시

[그림 1]은 주식 경향 예측에서 얻을 수 있는 세 가지 결과를 보여준다. 이 결과들은 서로 다른 예측 기법을 적용했을 때의 결과로 각 기법의 성능을 비교한다고 가정해보자. 정확도만 고려한다면 Result3가 가장 높은

성능을 보인다고 할 수 있다. 그러나 *Result3*의 정확도 외의 측정값을 보면 재현율과 특이도의 차이, 그리고 정밀도와 음성예측도의 차이가 상대적으로 큰 것을 알 수 있다. 재현율이 높은 이유는 이 모델이 65%의 테스트 벡터를 '상승'으로 예측하여, '상승'으로 레이블된 데이터 벡터들의 80%가 맞게 예측이 되었기 때문이다. 그러나 35%의 테스트 벡터만이 '하강'으로 예측되었기에 낮은 특이도를 가진다. 이렇게 재현율과 특이도의 차이가 큰 것은 예측이 균형적이지 않고 한쪽으로 치우쳤을 때 발생하며, 이런 결과를 내는 예측 모델은 좋은 모델이라고 할 수 없다. 그러나 정확도만 가지고 또는 일부의 측정값 (이 예에서는 재현율, 정밀도, 음성예측도)만을 가지고 예측 성능을 평가한다면 이 모델이 좋은 모델로 선정될 수 있다. 따라서 정확도 또는 그림에서 사용된 측정기준 중 일부분을 가지고 모델의 성능을 평가하는 것은 적절하지 않다.

*Result2* 역시 *Result1*보다 정확도는 약간 앞서지만, 나머지 측정값들을 고려하면 재현율과 특이도의 차이가 커서 예측이 편향되어 있음을 알 수 있다. 이 역시 정확도만 가지고 예측의 성능을 평가하는 것은 적절치 않고 정확도, 재현율, 특이도, 정밀도, 음성예측도 모두를 고려해야 함을 보여준다.

*Result1*의 정확도는 세 개의 결과 중 제일 낮으므로 정확도만 고려한다면 가장 안 좋은 평가를 받을 수밖에 없지만, 모든 측정값을 고려한다면 상대적으로 가장 균형적인 예측을 하고 있음을 알 수 있다. 따라서 [그림 1]을 통하여 기존의 연구와 같이 정확도만 가지고 또는 일부의 측정값만 가지고 예측의 성능을 평가하는 것은 예측의 성능을 온전히 나타낼 수 없으며 정확도, 재현율, 특이도, 정밀도, 음성예측도를 모두 고려해야 함을 알 수 있다. 본 절에서 주장하고자 하는 것은 상기 모든 측정기준을 사용하여 예측 성능을 평가해야 한다는 것이며, 구체적으로 여러 측정기준을 어떻게 이용하여 평가해야 하는지는 본 장의 4절에서 다루기로 한다.

[그림 2]는 KEPCO 주식을 가지고 얻은 결과로 SVM의 하이퍼-파라미터의 선택을 위해 여러 가지 하이퍼-파라미터 값을 시도하여 얻은 결과 중 일부이다. SVM 예측 모델의 성능은 그 모델에 관계된 하이퍼-파라미터에 따라 크게 달라질 수 있으며 가장 좋은 예측 성능을

보이는 하이퍼-파라미터를 선택하여 예측 모델을 학습해야 한다.

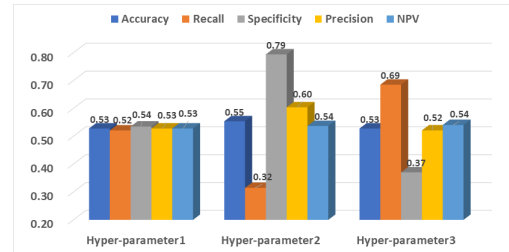


그림 2. 하이퍼 파라미터 선택 예시

세 가지의 하이퍼-파라미터 중 하나를 선택할 때 정확도만을 고려한다면 *Hyper-parameter2*를 선택해야 할 것이다. 그러나 다른 측정값들을 살펴보면 이 파라미터로 도출된 예측 모델은 재현율과 특이도의 차이가 굉장히 큰 편향된 모델이므로 선택에서 제외가 되어야 한다. 왜냐하면, 이 모델은 테스트 벡터의 상당 부분을 '하강'으로 예측했기 때문에 재현율이 낮고, 특이도가 높게 나온 것이기 때문이다. *Hyper-parameter3*는 모든 측정값을 주의 깊게 살펴보지 않으면 특이도를 제외한 다른 측정값들은 특별히 낮지 않기 때문에 문제가 없어 보인다. 그러나 모든 측정값을 잘 살펴보면 재현율은 높고 특이도는 상당히 낮은 것을 알 수 있다. 이러한 특징은 *Hyper-parameter3*로 만들어진 예측 모델이 상당 부분의 테스트 벡터를 '상승'으로 예측하기 때문에 발생한다. 반면에 *Hyper-parameter1*으로 학습된 모델의 모든 측정값들은 비슷한 값을 가진다. 이 특징은 '상승'으로 예측한 벡터의 개수와 '하강'으로 예측한 벡터의 개수가 비슷할 때 얻어진다.

본 절의 분석 결과를 정리하면, 추가 경향 예측 알고리즘을 비교하기 위한 측정기준으로 정확도뿐 아니라 재현율, 특이도, 정밀도, 음성예측도를 모두 제시해야 해당 알고리즘의 성능을 온전히 나타낼 수 있다. 또한, 하이퍼-파라미터 선정 시에도 상기 모든 측정기준을 고려해야 각 하이퍼-파라미터가 적용되어 만들어진 예측 모델의 성능을 정확하게 파악하여 최적의 하이퍼-파라미터를 선택할 수 있다.

### 3. 데이터 구성비율 분석

기계학습 기법을 사용하는 주식 경향 예측에 있어서 데이터는 많은 경우 학습 데이터 (training data), 검증 데이터 (validation data), 그리고 테스트 데이터 (test data)로 나누어진다. 학습 데이터는 모델을 만드는 데 사용되는 데이터이고, 검증 데이터는 학습되고 있는 모델을 위한 최적의 하이퍼-파라미터 (hyper-parameter)를 얻는 데 사용된다. 그리고 테스트 데이터는 학습된 모델의 성능을 평가하기 위해 사용된다. 이 세 가지의 데이터들은 주가의 '상승'과 '하강'으로 레이블된 벡터들을 포함하고 있는데 본 절의 분석 대상은 테스트와 검증 데이터 내에서 '상승'과 '하강'으로 레이블된 벡터들의 비율이다.

테스트와 검증 데이터 내의 '상승'으로 레이블된 벡터와 '하강'으로 레이블된 벡터의 비율은 하이퍼-파라미터 결정과 연구 결과의 공정한 비교라는 측면에서 중요하게 고려되어야 할 사항이지만 지금까지는 테스트와 검증 데이터의 구성비율에 대해 분석한 연구는 없었으며, 실험에 사용되는 데이터의 구성비율에 대해 언급하지 않는 경우가 많았다. 대개의 경우 전체 데이터를 어떤 비율로 학습, 검증, 테스트 데이터로 나누었는지 정도만 기술되고 있다. 따라서 대부분의 연구에서 사용된 '상승'으로 레이블된 벡터와 '하강'으로 레이블된 벡터의 비율을 알 수 없으며 관련 연구들이 따르는 어떠한 기준도 없는 것으로 보인다. 그러므로 본 절에서는 테스트와 검증 데이터 내의 '상승'으로 레이블된 벡터와 '하강'으로 레이블된 벡터의 비율이 하이퍼-파라미터 조정과 연구 결과의 비교에 어떻게 영향을 미치는지 분석하고자 한다.

#### 3.1 하이퍼-파라미터 조정에 미치는 영향

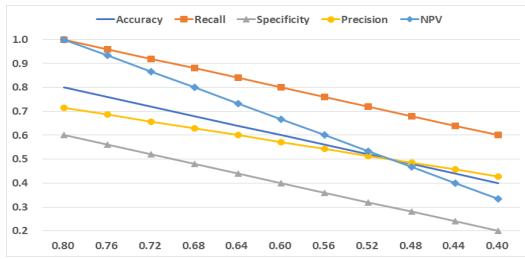
하이퍼-파라미터 (hyper-parameter)란 신경회로망 (neural network)이나 서포트벡터머신 (support vector machine)을 학습시킬 때 필요한 파라미터로 신경회로망의 경우 은닉층 (hidden layer)의 개수, 은닉층 뉴런의 개수, 학습률, 활성화함수 (activation function)의 종류 및 관련 파라미터 등을 의미하며, 서포트벡터머신의 경우 커널함수 (kernel function)의

종류 및 관련 파라미터, 그리고 코스트 파라미터 등을 포함한다. 하이퍼-파라미터의 설정에 따라 학습의 결과로 도출되는 예측 모델의 성능이 크게 달라지기 때문에 가장 좋은 성능을 보이는 예측 모델을 위한 하이퍼-파라미터를 찾는 것은 학습에서 매우 중요한 일이라 할 수 있다. 예측 모델을 위한 최적의 하이퍼-파라미터를 찾는 위해서는 일반적으로 하이퍼-파라미터를 변화시키며 예측 성능의 변화를 살펴보는 grid search 과정을 거치게 된다. 따라서 하이퍼-파라미터를 변화시킬 때 성능을 나타내는 측정값들이 가질 수 있는 값에 제한이 없어야 하며, 이상적인 값을 가질 수 있도록 보장되어야 한다.

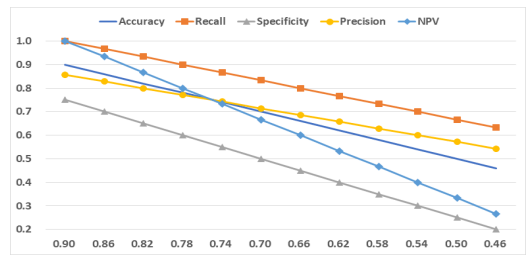
[그림 3]은 하이퍼-파라미터의 조정을 위한 검증 데이터 내에서 '상승'과 '하강'으로 레이블된 벡터들의 비율이 각각 50대50 그리고 60대40일 때, 이 두 가지 구성비율에 대해 다섯 가지의 '상승'과 '하강'으로 예측된 비율 (70대30, 60대40, 50대50, 40대60, 30대70)을 가정하여 정확도, 재현율, 특이도, 정밀도, 음성예측도가 가질 수 있는 값들을 보여준다. 그림에서 x축은 정확도의 값을 나타낸다.

'상승'과 '하강'으로 레이블된 벡터들의 비율이 50대50인 경우, 그리고 '상승'과 '하강'으로 예측된 비율이 50대50일 때([그림 3.(c)]), 정확도, 재현율, 특이도, 정밀도, 음성예측도는 동일한 값을 가진다. 그리고 '상승'과 '하강'으로 예측된 비율이 50대50에서 멀어질수록 재현율과 특이도의 차이가 더욱 벌어진다. 그리고 정밀도와 음성예측도의 차이는 재현율과 특이도의 차이보다 항상 작거나 같다.

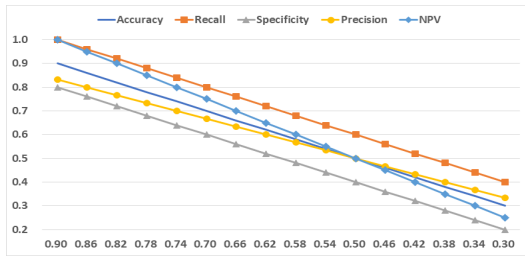
'상승'과 '하강'으로 레이블된 벡터들의 비율이 60대40인 경우, '상승'과 '하강'으로 예측된 비율이 60대40일 때([그림 3.(g)]), 재현율과 정밀도가 같고, 특이도와 음성예측도가 같다. 그리고 예측된 비율이 40대60일 때 ([그림 3.(i)]), 재현율과 음성예측도가 같고, 특이도와 정밀도가 같다. 예측된 비율이 더 차이가 크게 나는 방향으로 변하면, 재현율과 특이도의 차이와 정밀도와 음성예측도의 차이는 더욱 벌어진다.



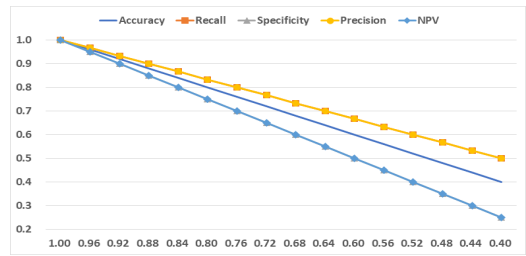
(a) 레이블된 비율: 50대50, 예측된 비율: 70:30 (PDR=0.4)



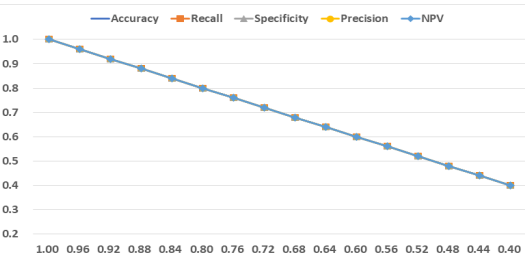
(f) 레이블된 비율: 60대40, 예측된 비율: 70:30



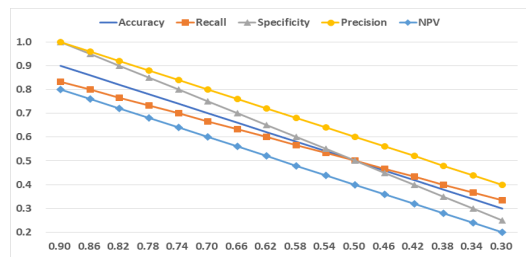
(b) 레이블된 비율: 50대50, 예측된 비율: 60:40 (PDR=0.2)



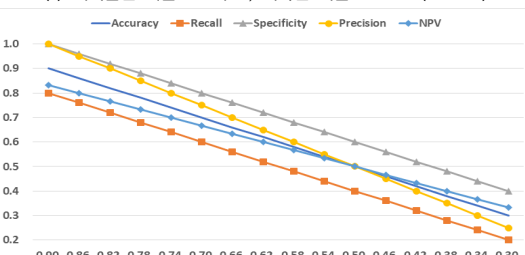
(q) 레이블된 비율: 60대40, 예측된 비율: 60:40



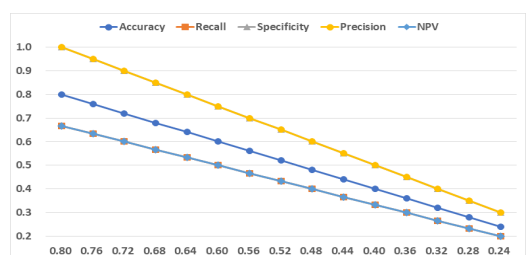
(c) 레이블된 비율: 50대50, 예측된 비율: 50:50 (PDR=0)



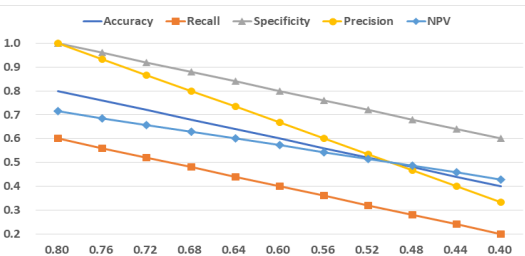
(h) 레이블된 비율: 60대40, 예측된 비율: 50:50



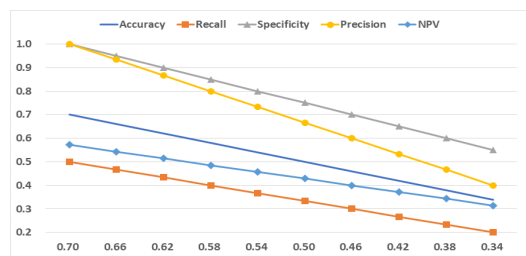
(d) 레이블된 비율: 50대50, 예측된 비율: 40:60 (PDR=0.2)



(i) 레이블된 비율: 60대40, 예측된 비율: 40:60



(e) 레이블된 비율: 50대50, 예측된 비율: 30:70 (PDR=0.4)



(j) 레이블된 비율: 60대40, 예측된 비율: 30:70

그림 3. '상승'과 '하강'으로 레이블된 비율, 예측된 비율에 따른 측정값의 변화

그림에서 알 수 있는 가장 중요한 것은 '상승'과 '하강'으로 레이블 된 벡터의 비율이 50대50이며 '상승'과 '하강'으로 예측된 비율이 50대50인 경우, 다섯 가지의 측정값이 모두 같아진다는 것이다. 이렇게 다섯 가지 측정값이 모두 같을 때가 측정기준 간에 균형이 제일 잘 맞는 경우라 할 수 있다. 만약 '상승'으로 예측된 벡터의 개수가 증가하면 재현율은 다른 측정기준들보다 향상될 수 있지만, '상승'으로 잘못 예측된 '하강' 벡터들이 늘어나 특이도와 정밀도는 감소할 수 있어 측정기준 간에 균형이 깨지게 된다. 이렇게 한 측정값이 향상되면 다른 측정값이 하강될 수 있으며, 따라서 다섯 가지의 측정값이 모두 같으면서 가장 높은 경우를 찾는 것이 가장 이상적이며 하이퍼-파라미터 조정의 목표라 할 수 있다.

그러나 그림에서 볼 수 있듯이 '상승'과 '하강'으로 레이블된 벡터의 비율이 60대40일 때는 이런 이상적인 경우는 얻을 수 없다. 이는 검증 데이터 내에서 '상승'으로 레이블된 벡터와 '하강'으로 레이블된 벡터의 개수가 같지 않으면 다섯 가지의 측정값이 모두 같은 이상적인 경우가 나올 수 없다는 것을 의미한다. 그러므로 이 경우에는 다섯 측정값이 모두 동일한 균형적인 예측을 하이퍼-파라미터 최적화의 목표로 설정하기 어렵다. 따라서 하이퍼-파라미터의 조정을 통해 최적의 예측 모델을 찾기 위해서는 검증 데이터의 구성비율이 '상승'으로 레이블된 벡터 50% 그리고 '하강'으로 레이블된 벡터 50%로 구성이 되는 것이 유리하다. 그리고 검증과 테스트 데이터의 구성을 '상승' 50%, '하강' 50%로 맞추는 것은 전처리과정에서 할 수 있다. 이 작업으로 인해 제거되는 벡터들이 전체 데이터 내의 다른 벡터들이 가지지 못한 어떤 패턴을 가지고 있을 수 있지만, 이것은 전체 데이터를 충분히 확보한다면 큰 문제가 되지 않는다. 또한, 이 작업은 모든 벡터들이 만들어진 후에 수행되므로 제거가 되지 않는 벡터들에게 주는 영향은 없다.

### 3.2 연구결과 제시에 관한 고찰

주가 경향 예측에 대한 새로운 알고리즘을 개발했다면 그 알고리즘을 검증하기 위하여 주가 데이터를 가지고 경향 예측의 성능을 평가하여 그 결과를 제시하게

된다. 이때, 실험 설정에 대한 정보도 함께 제공되어야 하는데, 이 정보에는 검증 데이터와 테스트 데이터의 '상승'과 '하강' 벡터의 구성비율도 포함되어야 한다. [그림 3]에서 볼 수 있듯이 데이터의 구성이 달라지면 같은 모델이라도 얻을 수 있는 정확도, 재현율 등과 같은 측정값이 달라지기 때문에, 그 연구를 보고 결과를 재현하거나, 다른 연구 결과와 비교하기 위해서는 데이터 구성에 대한 정보가 필요하다. 또한, 동일한 데이터 구성을 사용하여 연구결과를 도출한다면, 즉 연구자들 사이에 어떤 기준이 있다면, 여러 주가 경향 예측 알고리즘의 성능을 더욱 쉽게 비교할 수 있을 것이다.

그렇다면 어떤 데이터 구성을 사용하여 연구 결과를 얻는 것이 타당하며 기준이 될 수 있을까? 이전 절에서 하이퍼-파라미터의 최적화를 위해서 50대50의 데이터 검증 데이터의 구성을 사용하는 것이 필요하다고 하였다. 따라서 하이퍼-파라미터의 최적화를 위해 사용했던 50대50의 비율을 테스트 데이터의 구성에도 사용하는 것이 예측 모델의 성능을 온전히 나타낼 수 있을 것이다. 예를 들어 50대50의 데이터 구성을 가지고 하이퍼-파라미터를 최적화하여 예측 모델을 만들었고, 이 모델의 검증 데이터에 대한 정확도, 재현율, 특이도 등의 측정값들이 서로 비슷한 수준이었으며, 50대50으로 구성된 테스트 데이터에 대한 측정값들도 서로 비슷하다고 가정하자. 그런데 테스트 데이터의 구성에 60대40의 비율을 사용하여 동일 모델의 성능을 검증하면 [그림 3]에서처럼 60대40으로 구성된 데이터에 대한 측정값들이 가질 수 있는 범위가 50대50으로 구성된 데이터에 대한 측정값의 범위와 달라서 상대적으로 예측의 균형이 떨어진 측정결과가 도출될 것이다. 즉 도출된 결과가 모델의 성능을 온전히 나타내지 못한다. 따라서 주가 경향 예측 연구에 있어서 예측 성능 측정에 사용된 데이터의 구성비율은 제시되어야 하며 그 비율은 50대50으로 하는 것이 가장 적절하다.

### 4. 주가 경향 예측 성능의 측정기준 제안

본 장의 2절에서는 주가 예측 모델의 성능을 온전히 나타내기 위해서 정확도, 재현율, 특이도, 정밀도, 음성 예측도가 모두 사용되어야 한다고 하였다. 그렇다면 두 개의 예측 모델을 비교할 때 이 다섯 가지의 측정값을



어떻게 이용해야 하는가? 다섯 가지의 측정값을 모두 고려하여 두 모델을 비교할 때, 한 모델의 다섯 측정값이 모두 다른 모델의 측정값보다 높거나 낮지 않다면 어느 모델이 우월한지 판단하기 쉽지 않다. 그렇다면 비교를 하기 위해 다섯 가지의 측정값을 하나의 값으로 표현하는 방법을 사용할 수 있을 것이다. 예를 들어 F-score가 재현율과 정밀도를 조화평균 (harmonic mean)으로 합성한 것과 같이 다섯 가지 측정값의 평균을 이용할 수 있다.

표 3. 평균을 새로운 측정기준으로 사용한 예시

	Model1	Model2	Model3
정확도	0.680	0.680	0.690
재현율	0.680	0.730	0.840
특이도	0.680	0.630	0.540
정밀도	0.680	0.664	0.646
음성예측도	0.680	0.700	0.771
산술평균	0.680	0.681	0.695
조화평균	0.680	0.679	0.682
기하평균	0.680	0.680	0.690

[표 3]은 세 모델의 다섯 가지 측정값과 세 가지 평균값 (산술평균, 조화평균, 기하평균)을 보여준다. Model1은 모든 측정값들이 같은 경우로 가장 이상적인 경우이며 조화평균 역시 측정값들과 동일하다. 이 모델과 Model2를 비교해 보면 Model2는 Model1과 같은 정확도를 가졌지만 다른 측정값들 사이에는 차이가 존재한다. 이 두 모델의 평균값들을 비교하면 거의 같은 값을 가진다. 그러나 Model2의 재현율과 특이도, 그리고 정밀도와 음성예측도 간의 차이에도 불구하고 거의 같은 값을 가진다는 것은 평균값들이 측정값 간의 차이를 잘 반영하지 못함을 나타낸다. Model1과 Model3를 비교하면 Model3는 Model1보다 정확도가 1.5% 높지만 다른 측정값 간의 차이는 상당히 심하다. 그럼에도 불구하고 Model3의 평균값들은 모두 Model1보다 높다. 그러나 아무도 Model3를 Model1보다 선호하지 않을 것이며, 역시 평균값을 가지고 모델의 좋고 나쁨을 따지는 것은 부적절하다는 것을 보여준다. 그러므로 하나의 값으로 모델의 성능을 제대로 표현할 수 없으며 정확도에 측정값 간의 차이의 정도를 나타내는 지표를 추가하여 사용하는 것이 적절할 것이다. 여기서 측정값 간의 차이는 재현율과 특이도의 차

이, 그리고 정밀도와 음성예측도의 차이를 의미하며 이들 간의 차이가 줄어들수록 모든 다섯 개의 측정값들로 비슷해지게 된다.

그렇다면 어떻게 측정값 간의 차이의 정도를 나타낼 수 있는가? 만약 3절에서 분석한 대로 데이터의 구성비율, 즉 '상승'과 '하강'으로 레이블된 벡터들의 비율을 50대50으로 맞춘다면 '상승'과 '하강'으로 예측된 비율이 측정값 간의 차이를 나타낼 수 있다. [그림 3.(a)-(e)]에서 볼 수 있듯이 '상승'과 '하강'으로 레이블된 벡터들의 비율이 동일한 경우의 특징은 '상승'과 '하강'으로 예측된 비율이 50대50에서 멀어질수록 재현율과 특이도의 차이가 커지고, 정밀도와 음성예측도의 차이는 재현율과 특이도의 차이보다 항상 작거나 같다는 것이다. 즉 '상승'과 '하강'으로 예측한 비율로 측정값 간의 불균형 정도를 나타낼 수 있으며 정확도와 함께 모델의 예측 성능을 나타낼 수 있다. 이 정확도와 사용될 수 있는 지표값인 예측 불균형 비율 (prediction disparity ratio; PDR)을 정의하면 다음과 같다.

$$PDR = \frac{|NPP - NNP|}{NP} \quad (1)$$

여기서 NPP (the number of positive predictions)는 주가를 '상승'으로 예측한 횟수, NNP (the number of negative predictions)는 주가를 '하강'으로 예측한 횟수, 그리고 NP (the number of predictions)는 총 예측의 횟수를 의미한다. 만약 PDR이 1이라면 예측은 한쪽 방향으로만 이루어졌음을 의미하고, 만약 PDR이 0이라면 예측이 각 방향으로 동일하게 이루어졌음을 의미한다. 여기서 중요한 것은 PDR이 모델의 예측 불균형 정도를 나타내려면 그 전제 조건이 데이터 내의 '상승'으로 레이블된 벡터의 개수와 '하강'으로 레이블된 벡터들의 개수가 동일해야 한다는 것이다. 이 조건이 맞지 않으면 PDR이 예측 불균형의 정도를 정확히 나타내지 못한다. 왜냐하면, 예를 들어 '상승'으로 레이블된 데이터가 많다면 자연스럽게 '상승'으로 예측된 데이터도 많을 것인데, 식 (1)에서는 그것을 불균형으로 계산하기 때문이다.

PDR은 '상승'으로 레이블된 벡터의 개수와 '하강'으로 레이블된 벡터의 개수가 같다는 조건이 만족된다면,

재현율과 특이도의 차이와 같은 값을 가지며, 정밀도와 음성예측도의 차이는 재현율과 특이도의 차이보다 작거나 같으므로 정확도와 함께 예측의 성능을 표현하기에 충분하다. [그림 3.(a)-(e)]에 표시된 PDR을 살펴보면, PDR이 재현율과 특이도의 차이와 같은 것을 알 수 있으며, 정밀도와 음성예측도의 차이는 PDR보다 항상 작고, 정확도는 재현율과 음성예측도의 중간값인 것을 알 수 있다. 예측 모델이 어떤 비율로 '상승'과 '하강'으로 예측하는지에 따라 PDR은 달라지겠지만, PDR과 측정값들과의 관계는 변하지 않으며, 이 특징 때문에 PDR이 예측 모델이 얼마나 균형있는 예측을 수행하는지 나타낼 수 있다.

만약 측정값들 간의 차이의 평균을 지표로 사용하고 해도, 이 지표값은 평균값이므로 어떤 측정값 간의 차이는 크고, 다른 차이는 작을 수 있다. 반면에 PDR은 측정값 간의 차이에 대한 좀더 확실한 정보를 제공함과 동시에 계산이 더 간단하다. 작은 PDR을 가진다는 것은 작은 측정값 간의 차이를 보장하기 때문에, 어떤 예측 모델이 어느 정도 값 이하의 작은 PDR을 가진다면, 그 모델의 정확도만으로 그 모델의 성능을 나타낼 수 있다. 예를 들어 하이퍼-파라미터를 최적화할 때, 어떤 범위 내의 PDR을 가지는 예측 모델 중 가장 높은 정확도를 가지는 예측 모델을 선택하여, 그 예측 모델을 만들 때 사용된 하이퍼-파라미터를 최적의 하이퍼-파라미터로 선택할 수 있다.

### 5. 예측 모델의 성능 평가 방법에 대한 제안

앞 절의 분석을 바탕으로 다음과 같은 주가 경향 예측 모델의 성능 평가 방법을 제안한다.

- 테스트 데이터와 검증 데이터를 구성할 때 '상승'으로 레이블된 벡터의 개수와 '하강'으로 레이블된 벡터의 개수를 같게 맞춘다.
- 예측 모델의 성능을 나타낼 때, 그 모델의 성능은 예측의 정확성을 나타내는 정확도와 예측의 균형을 나타내는 PDR로 표현할 수 있다.
- 2개 이상의 예측 모델 중 가장 성능이 좋은 모델을 선택해야 할 때에는 좋은 예측 균형을 보장하는 어떤 범위 내의 PDR을 가지는 모델 중 가장 정확도가 높은 모델을 선택한다.

이 성능 평가 방법의 장점은 PDR을 이용하여 재현율, 특이도, 정밀도, 음성예측도를 모두 고려하는 효과를 가지면서, 간단하게 정확도만으로 성능을 나타낼 수 있다는 것이다. 따라서 하이퍼-파라미터의 최적화와 다른 연구결과와의 비교를 온전하게, 그리고 쉽게 수행할 수 있다.

그럼 실제로 학습된 모델이 보여주는 PDR과 다섯 가지 측정값의 관계를 살펴보도록 하자. [표 4]는 각 주식에 대해 하이퍼-파라미터를 조절하여 가장 높은 정확도를 얻은 경우 (*Highest acc.*), 가장 높은 정확도와와의 차이가 5% 이내인 정확도 중 가장 낮은 PDR을 얻은 경우 (*Lowest PDR within top 5% acc.*), 가장 높은 정확도와와의 차이가 5% 이내인 정확도 중 가장 높은 PDR을 얻은 경우 (*Highest PDR within top 5% acc.*), PDR이 0.1 이하인 경우 중에 가장 높은 정확도를 얻은 경우 (*Highest acc. with PDR of 0.1*)에 대하여 정확도 (accuracy), 재현율 (recall), 특이도 (specificity), 정밀도 (precision), 음성예측도 (NPV)를 정리하였다. 표에서 보여주는 주식별 네 가지 경우는 하이퍼-파라미터를 바꿔가며 만든 다수의 예측 모델 중 특별히 상기 조건에 맞게 선택된 네 개 예측 모델로부터 얻어지는 결과이다. 하이퍼-파라미터는 모델을 나타내는 수식에서 사용되는 파라미터로 파라미터의 값에 따라 모델의 성능, 즉 정확도 등이 달라진다. 즉 표의 결과는 주식별로 동일한 학습 데이터와 검증 데이터를 사용하여 얻어지는 네 개의 SVM 예측 모델로부터의 결과이며, 동일 주식의 모델끼리 다른 유일한 점은 하이퍼-파라미터의 값이다.

가장 높은 정확도를 얻은 경우, 모든 주식에 대하여 PDR이 0.14보다 컸으며 이는 가장 높은 정확도와 가장 낮은 PDR은 동시에 얻기 힘들다는 것을 뜻한다. 특히 KEPCO의 경우는 PDR이 0.479일 때 최고의 정확도가 얻어지며 이 경우 재현율과 특이도의 차이가 커서 모델이 한쪽으로 많이 편향된 예측을 하고 있음을 알 수 있다. 따라서 정확도는 가장 높게 나오지만 제대로 된 모델이라 할 수 없다.

가장 높은 정확도의 5% 이내의 정확도 중 가장 낮은 PDR을 얻은 경우를 보면 재현율과 특이도의 차이가 Mobis를 제외하고 0.03보다 작으며 Mobis의 경우에

표 4. 다양한 모델의 측정값 및 PDR 비교

KOSPI200						
	Accuracy	Recall	Specificity	Precision	NPV	PDR
Highest acc.	0.569	0.493	0.644	0.581	0.560	0.151
Lowest PDR within top 5% acc.	0.562	0.563	0.561	0.562	0.562	0.003
Highest PDR within top 5% acc.	0.542	0.221	0.863	0.617	0.526	0.642
Highest acc. within PDR of 0.1	0.562	0.563	0.561	0.562	0.562	0.003
Samsung Electronics						
	Accuracy	Recall	Specificity	Precision	NPV	PDR
Highest acc.	0.615	0.545	0.686	0.634	0.601	0.142
Lowest PDR within top 5% acc.	0.592	0.579	0.606	0.595	0.590	0.028
Highest PDR within top 5% acc.	0.611	0.302	0.920	0.790	0.568	0.618
Highest acc. within PDR of 0.1	0.609	0.566	0.652	0.620	0.601	0.086
Industry bank						
	Accuracy	Recall	Specificity	Precision	NPV	PDR
Highest acc.	0.563	0.467	0.658	0.578	0.553	0.191
Lowest PDR within top 5% acc.	0.538	0.533	0.542	0.538	0.537	0.009
Highest PDR within top 5% acc.	0.549	0.241	0.856	0.626	0.530	0.614
Highest acc. within PDR of 0.1	0.549	0.508	0.589	0.553	0.545	0.082
KEPCO						
	Accuracy	Recall	Specificity	Precision	NPV	PDR
Highest acc.	0.554	0.315	0.794	0.604	0.537	0.479
Lowest PDR within top 5% acc.	0.528	0.521	0.535	0.528	0.528	0.014
Highest PDR within top 5% acc.	0.549	0.248	0.850	0.622	0.530	0.602
Highest acc. within PDR of 0.1	0.543	0.507	0.579	0.547	0.540	0.072
Mobis						
	Accuracy	Recall	Specificity	Precision	NPV	PDR
Highest acc.	0.551	0.478	0.624	0.560	0.545	0.146
Lowest PDR within top 5% acc.	0.544	0.497	0.591	0.549	0.540	0.094
Highest PDR within top 5% acc.	0.541	0.218	0.865	0.617	0.525	0.646
Highest acc. within PDR of 0.1	0.544	0.497	0.591	0.549	0.540	0.094

도 0.1이 되지 않아 균형 잡힌 예측을 하고 있지만, 정확도는 최고 정확도보다 평균 3.1% 낮다.

가장 높은 정확도의 5% 이내의 정확도 중 가장 높은 PDR을 얻은 경우를 살펴보면 최고 정확도보다 평균 2.1% 낮은 정확도이지만 PDR은 모두 0.6이 넘어 예측 모델로 선택되기에는 부적절하다. 가장 낮은 PDR을 가지는 경우와 비교하면, 정확도는 비슷하지만, 예측의 균형 측면에서는 많이 다를 수 있다는 것을 보여준다.

마지막으로 PDR이 0.1보다 작은 경우 얻을 수 있는 최고의 정확도 (*highest acc. with PDR of 0.1*)를 보면 평균 0.067의 PDR을 가져 측정값 간의 차이가 크지 않으며, 이 때의 평균 정확도는 얻을 수 있는 최고 정확도보다 1.59% 떨어지는 것으로 분석되었다. 가장 높은 정확도를 얻을 때 PDR은 0.22인 것과 비교하면 보다 적절한 예측 모델이라고 할 수 있다. 따라서 예측 모델의 성능을 평가할 때 어느 정도의 PDR ([표 4]의 예에서는 0.1) 내에서 얻을 수 있는 최고의 정확도를 보

이는 모델을 선택하는 것이 가장 타당하다고 할 수 있으며 이 예측 모델의 평가 기준의 사용을 제안한다.

#### IV. 결론

주가 경향 예측에 있어서 예측의 성능을 평가하기 위한 측정기준으로는 정확도를 많이 사용해왔지만, 정확도만으로는 예측의 성능을 제대로 나타낼 수 없기 때문에 재현율, 정밀도, 특이도, 음성예측도 함께 고려되어야 한다. 또한 주가 경향 예측 연구에서 사용되는 테스트 데이터와 검증 데이터를 구성할 때, '상승'으로 레이블된 데이터와 '하강'으로 레이블된 데이터의 비율을 같게 맞추는 것은 하이퍼-파라미터의 최적화와 연구 결과의 공정한 제시의 측면에서 꼭 필요하다.

따라서 하이퍼-파라미터의 최적화와 연구결과와 공정한 제시를 위해서는 '상승'과 '하강' 데이터의 비율을

동일하게 맞추고, 성능 평가를 위한 측정기준으로 정확도 뿐 아니라 재현율, 정밀도, 특이도, 음성예측도 함께 고려해야 한다. 하지만 이 모두를 고려하여 예측의 성능을 평가하기는 쉽지 않다. 이에 본 논문에서는 주가 경향 예측의 성능을 평가하기 위해 예측 불균형 비율(PDR)에 기반을 둔 방법을 제시하였다. 이 성능 평가 방법은 재현율, 특이도, 정밀도, 음성예측도를 모두 표현하는 PDR과 정확도만으로 성능을 나타낼 수 있다. 이 방법이 주식 경향 예측 연구에 널리 적용된다면 최적의 예측 모델을 만들 수 있고, 연구 결과의 공정하고 용이한 비교가 가능할 것으로 기대된다.

본 논문은 주가의 경향을 '상승'과 '하강'로 예측하는 이진 분류를 대상으로 하고 있다. 그러나 주가의 경향을 세 가지 이상으로 분류하는 연구도 있으므로 이런 경우의 성능 평가 방법에 관한 연구를 진행할 계획이다. 또한 본 논문에서는 성능의 평가 기준으로 정확도를 대상으로 했지만, 주가 예측 모델의 또 다른 성능 평가 기준인 수익률에 대한 연구도 진행할 예정이다.

**참 고 문 헌**

[1] H. S. Sim, H. I. Kim, and J. J. Ahn, "Is deep learning for image recognition applicable to stock market prediction?," Complexity, Vol.2019, pp.1-10, 2019.  
 [2] <https://www.index.go.kr/main.do>  
 [3] X. Zhang, Y. Hu, K. Xie, S. Wang, E. Ngai, and M. Liu, "A causal feature selection algorithm for stock prediction modeling," Neurocomputing, Vol.142, pp.48-59, 2014.  
 [4] <https://en.wikipedia.org/wiki/Robo-advisor>  
 [5] <http://www.ratested.kr/portal/main/main.do>  
 [6] E. Hoseinzade and S. Haratizadeh, "CNNpred: CNN-based stock market prediction using a diverse set of variables," Expert Systems with Applications, Vol.129, pp.273-285, 2019.  
 [7] W. Chiang, D. ENke, T. Wu, and R. Wang, "An adaptive stock index trading decision support system," Expert Systems with Applications, Vol.59, pp.195-207, 2016.

[8] J. Zhang, S. Cui, Y. Xu, Q. Li, and T. Li, "A novel data-driven stock price trend prediction system," Expert Systems with Applications, Vol.97, pp.60-69, 2018.  
 [9] Y. Chen and Y. Hao, "Integrating principle component analysis and weighted support vector machine for stock trading signals prediction," Neurocomputing, Vol.321, pp.381-402, 2018.  
 [10] C. Tsai and Y. Hsiao, "Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches," Decision Support Systems, Vol.50, pp.258-269, 2010.  
 [11] Z. Hu, J. Zhu, and K. Tse, "Stocks market prediction using support vector machine," Int. Conf. on Information Management, Innovation Management and Industrial Engineering, pp.115-118, 2013.  
 [12] J. Patel, S. Shah, P. Thakkar, and K. Kotecha, "Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques," Expert Systems with Applications, Vol.42, pp.259-268, 2015.  
 [13] A. N. Kia, S. Haratizadeh, and S. B. Shouraki, "A hybrid supervised semi-supervised graph-based model to predict one-day ahead movement of global stock markets and commodity prices," Expert Systems with Applications, Vol.105, pp.159-173, 2018.  
 [14] W. Long, Z. Lu, and L. Cui, "Deep learning-based feature engineering for stock price movement prediction," Knowledge-Based Systems, Vol.164, pp.163-173, 2019.  
 [15] 박강희, 신현정, "Semi-supervised learning을 이용한 주가예측," 한국경영과학회 학술대회 논문집, pp.110-116, 2010.  
 [16] 안성원, 조성배, "뉴스 텍스트 마이닝과 시계열 분석을 이용한 주가예측," 한국정보과학회 학술발표 논문집, pp.364-369, 2010.  
 [17] 신동근, 정경용, "웨이블릿 변환과 퍼지 신경망을 이용한 단기 KOSPI 예측," 한국콘텐츠학회논문지, 제

- 11권, 제6호, pp.1-7, 2011.
- [18] 김진화, 홍광현, 민진영, “지식 누적을 이용한 실시간 주식시장 예측,” 지능정보연구, 제17권, 제4호, pp.109-130, 2011.
- [19] 허준영, 양진용, “SVM 기반의 재무 정보를 이용한 주가 예측,” 정보과학회 컴퓨팅의 실제 논문지, 제21권, 제3호, pp.167-172, 2015.
- [20] Z. Lei and W. Lin, “Price trend prediction of stock market using outlier data mining algorithm,” IEEE Int. Conf. on Big Data and Cloud Computing, pp.93-98, 2015.
- [21] G. Dong, K. Fataliyev, and L. Wang, “One-step and multi-step ahead stock prediction using backpropagation neural networks,” Int. Conf. on Information, Communication, and Signal Processing, pp.1-5, 2013.
- [22] F. Wang, Z. Zhao, X. Li, and H. Zhang, “Stock volatility prediction using multi-kernel learning based extreme learning machine,” Int. Joint Conf. on Neural Networks, pp.3078-3085, 2014.
- [23] Y. Xu, Z. Li, and L. Luo, “A study on feature selection for the trend prediction of stock trading price,” Int. Conf. on Computational and Information Sciences, pp.579-582, 2013.
- [24] K. Kim, “Financial time series forecasting using support vector machine,” Neurocomputing, Vol.55, pp.307-319, 2003.
- [25] Y. Lin, H. Guo, and J. Hu, “An SVM-based approach for stock market trend prediction,” Int. Joint Conf. on Neural Networks, 2013.
- [26] A. K. Sirohi, P. K. Mahato, and V. Attar, “Multiple kernel learning for stock price direction prediction,” IEEE Int. Conf. on Advances in Engineering & Technology Research. 2014.
- [27] D. Kato and T. Nagao, “Stock prediction using multiple time series of stock prices and news articles,” IEEE Symposium on Computers & Informatics, pp.11-16, 2012.
- [28] Y. Luo, J. Hu, X. Wei, D. Fang, and H. Shao, “Stock trends prediction based on hypergraph modeling clustering algorithm,” IEEE Int. Conf. on Progress in Informatics and Computing, pp.27-31, 2014.
- [29] K. Kim and I. Han, “Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index,” Expert Systems with Applications, Vol.19, pp.125-132, 2000.
- [30] A. Oztekin, R. Kizilaslan, S. Freund, and A. Iseri, “A data analytic approach to forecasting daily stock returns in an emerging market,” European Journal of Operational Research, Vol.253, pp.697-710, 2016.
- [31] M. Lee, “Using support vector machine with hybrid feature selection method to the stock trend prediction,” Expert Systems with Applications, Vol.36, pp.10896-10904, 2009.
- [32] L. Ni, Z. Ni, and Y. Gao, “Stock trend prediction based on fractal feature selection and support vector machine,” Expert System with Applications, Vol.38, pp.5569-5576, 2011.
- [33] D. Kumar, S. S. Meghwani, and M. Thakur, “Proximal support vector machine based hybrid prediction models for trend forecasting in financial markets,” Journal of Computational Science, Vol.17, pp.1-13, 2016.

#### 저 자 소 개

#### 임 정 수(Chungsoo Lim)

#### 정회원



- 2009년 5월 : North Carolina State University, Computer Engineering(공학박사)
- 2013년 10월 ~ 현재 : 한국교통대학교 전자공학과 부교수

<관심분야> : 주가 경향 예측, 생체 인식, 기계학습 가속기