



PCA를 활용한 기업실적 예측변수 생성

Generating Firm's Performance Indicators by Applying PCA

이준혁* · 김갑조* · 박상상** · 장동식*[†]

Joonhyuck Lee*, Gabjo Kim*, Sangsung Park**, and Dongsik Jang*[†]

*고려대학교 산업경영공학과, **고려대학교 지식재산학과

*Department of Industrial Management Engineering, Korea University

**Department of Intellectual Property, Korea University

요약

최근 기업의 실적 및 주가를 예측하기 위해 매출액증가율, 부채비율 등의 다양한 예측변수를 활용하여 정량적인 예측방법을 활용하는 연구가 많이 이루어지고 있다. 기업실적 및 주가를 정량적 예측하기 위해 수많은 예측변수들 중에서 모델구축을 위해 중요한 예측변수를 선정하는 것이 중요하다. 대부분의 기존연구들에서는 다양한 알고리즘을 활용하여 예측변수들을 제거하는 방법을 사용하는 경우가 많았다. 이러한 경우 각 예측변수들이 가지는 많은 정보들이 제거되는 문제점이 존재한다. 이러한 문제점을 해결하기 위해 본 연구에서는 예측모델 구축을 위해 예측변수들을 제거하는 대신 각 변수들이 가지고 있는 정보를 병합하여 새로운 변수를 생성하는 대표적인 차원축소 방법인 주성분분석(PCA)을 활용하였다. 본 연구에서는 제안된 예측모델을 미국의 전자, 전기기업의 재무정보를 활용하여 구축하고 예측성능을 실증적으로 분석해 보았다.

키워드 : 유전알고리즘, 인공신경망, 주성분분석, 실적예측, 예측모델

Abstract

There have been many studies on statistical forecasting on firm's performance and stock price by applying various financial indicators such as debt ratio and sales growth rate. Selecting predictors for constructing a prediction model among the various financial indicators is very important for precise prediction. Most of the previous studies applied variable selection algorithms for selecting predictors. However, the variable selection algorithm is considered to be at risk of eliminating certain amount of information from the indicators that were excluded from model construction. Therefore, we propose a firm's performance prediction model which principal component analysis is applied instead of the variable selection algorithm, in order to reduce dimensionality of input variables of the prediction model. In this study, we constructed the proposed prediction model by using financial data of American IT companies to empirically analyze prediction performance of the model.

Key Words : Genetic Algorithm, Artificial Neural Network, Principal Component Analysis, Performance Prediction, Prediction Model.

Received: Sep. 14, 2014
Revised : Sep. 28, 2014
Accepted: Jan. 5, 2015
[†]Corresponding author(jang@korea.ac.kr)

1. 서론

기업 및 주식 투자 시 투자이익을 향상시키기 위해 기업의 실적과 그에 따른 주가변동을 예측하기 위한 연구가 많이 이루어졌다. 전통적인 기업실적 예측 관련연구들은 기업실적에 유의한 영향을 미치는 것으로 알려진 재무지표, 전기 주가정보 등의 다양한 정보를 종합하여 연구자의 정성적인 분석을 통해 이루어졌다. 하지만 이와 같이 연구자의 주관적 판단에만 의존한 연구들은 객관성이 부족하다는 한계점이 존재한다. 최근 들어 데이터 처리 및 기계학습 기술의 발달에 따라 대량의 데이터로부터 유의한 결과를 도출하고 이를 예측에 활용할 수 있게 되었다. 이에 따라 기업 실적 및 주가예측과 관련된 연구들도 전통적인 정성적 분석 방법 외에 인공신경망, 의사결정나무, SVM 등을 통해 정량적으로 기업실적을 예측하려는 시도가 많이 이루어지고 있다[1, 2, 3, 5]. 기계학습을 통해 정량적으로 기업의 실적을 예측하는 경우, 연구자가 기업실적예측 모델에 사용되는 예측변수를 선정하여야 한다. 예측모델 구축을 위한 예측변수 선정은 모델구축을 위한 학습알고리즘 선정 못지않게 중요하다. 기존연구들은 자산증가율, 자본증가율, 전년도 매출액 등과 같이 기업 실적 및 주가에 유의한 영향을 미치는 예측변수들을 도출하고 이들을 모델학습에 사용하였다[1,2,3]. 하지만 모델 구축에 너무 많은 수의 예측변수를 사용하면 변수들 간의 다중공선성과 같은 교호작용, 상관성 등의 문제로 인해 예측성능이 저하될 가능성이 존재한다. 이러한 문제점을 해결하기 위해 기존연구들에서는 SSVS(Stochastic Search Variable Selection), GVS(Gibbs Variable Selection)등의 다양한 알고리즘을 활용하여 모델구축에 적합한 예측변수들을 선정하였다[5]. 위와

이 논문은 BK21 플러스 사업(고려대학교, 제조·물류분야에서의 빅데이터 운용 사업팀)으로 지원된 연구임
이 논문은 2012년 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(한국연구재단-NRF-2010-0024163)
This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

같은 변수선택 알고리즘들은 다수의 예측변수들 중 소수의 변수들을 선정하는 방식이기 때문에 변수들을 제거하는 과정에서 유의미한 정보량을 손실할 가능성이 존재한다. 본 연구에서는 위와 같은 문제점을 해결하기 위해 다수의 예측변수들 중에서 특정 변수들을 선정하는 방식이 아닌 변수들이 가지고 있는 정보들을 최대한 확보하는 적은 수의 새로운 변수들을 생성하는 PCA(Principle Component Analysis)를 활용한 기업실적 예측모델을 제안한다[4]. 본 연구에서는 제안된 예측모델을 구축하기 위해 51개 미국의 전기, 전자기업들을 대상으로 12년간의 기업실적 정보 및 재무정보를 수집하였다. 이를 바탕으로 본 연구에서 제안한 예측모델을 구축하고 그 성능을 실증적으로 분석 하였다.

2. 이론적 배경

2.1 기존 연구

경영학 및 경제학 분야에서는 기업의 성과와 경기의 흐름을 예측하고자 하는 연구가 지속적으로 이루어졌다. 하지만 많은 기존연구들은 기업성과 및 경제상황 등을 예측하기 위해 정량적인 방법이 아닌 정성적인 평가방법에 많이 의존하였다. 하지만 최근 의사결정나무, 인공신경망, SVM 등과 같은 기계학습 알고리즘들이 많이 개발되면서 정량적으로 기업의 실적, 기업주가 등을 예측하고자 하는 연구가 활발히 진행되고 있다. 이와 관련된 대표적인 연구분야는 기업의 도산예측과 주가예측이 있다[1,2,3,5].

먼저 부도예측과 관련된 연구는 다양한 재무지표들 중 경영안정성에 관련된 변수들을 선별하여 이를 예측변수로 활용하고 이를 통해 기업의 도산여부를 예측하는 연구가 주를 이루었다. 이와 관련된 연구들에서는 로지스틱회귀분석, 인공신경망 등의 알고리즘이 주로 사용되었다[1,2,3].

기업 주가예측과 관련된 연구로는 다양한 예측지표를 활용하여 주가의 등락 또는 특정 회사의 주가를 예측하는 연구가 많이 이루어 졌다[5].

이와 같이 기업의 부도여부 및 주가의 등락을 예측하는 연구는 많이 이루어 졌지만 기업의 매출액, 순이익과 같이 기업의 실적을 예측하는 연구는 비교적 활발히 수행되지 않았다.

2.2 PCA

본 논문에서는 다양한 예측지표들 간의 상관성 및 교호작용 등의 문제점을 해결하기 위해 Dimension Reduction의 대표적인 기법인 PCA(Principial Component Analysis)를 활용하였다. PCA는 여러 변수들($x_1, x_2, x_3, \dots, x_n$)이 관측되었을 때, 상관성이 높은 변수들을 공동요인으로 묶어 기존 변수들이 가지고 있는 정보들을 최대한 확보하는 적은 수의 새로운 변수들을 생성하는 방법이다[4]. 이와 같은 PCA의 특성 때문에 다양한 방면의 텍스트 마이닝과 다변량 통계분석 관련 연구에서 널리 사용된다[6,7].

2.3 인공신경망

현재 기계학습을 활용한 알고리즘들 중 인공신경망(ANN)이 기업 실적 예측을 위한 연구에 널리 사용되고 있다. 인공 신경망은 사람의 두뇌가 반복적 경험을 통해 학습해가는 것을 모방한 알고리즘으로, 주어진 학습 데이터를 통해 반복적인 학습을 함으로써 특정한 패턴을 찾아내고 이를 통해 미래의 값을 예측한다. 인공신경망은 특히 독립변수와 종속변수 사이의 관계를 명확히 설명하기 어렵고 복잡한 경우

에도 비교적 좋은 결과를 제공한다는 장점이 있다. 인공신경망의 구조는 입력계층과 출력계층 그리고 은닉계층으로 이루어져 있다. 입력노드를 통해 들어온 독립변수를 은닉노드가 전달받아 선형결합을 한다. 예를 들어 X_1, X_2, \dots, X_n 개의 설명변수가 있을 때 은닉노드(HN)는 다음 식(1)와 같이 표현된다.

$$\text{은닉노드(HN)} = w_1X_1 + w_2X_2 + \dots + w_nX_n \quad (1)$$

인공신경망 모형에서 다수의 은닉노드가 존재하는 경우 가중치 w_n 은 노드들 마다 다르게 설정된다. 출력노드는 은닉노드들로부터 나온 값을 전달받아 결과 값을 도출한다. 신경망을 통해 예측하고자 하는 종속변수가 연속형인 경우에는 선형결합을 통해 결과 값을 도출하고, 범주형인 경우에는 결과 값을 범주형 값으로 변환하여 나타낸다.

2.4 유전알고리즘

유전알고리즘은 자연의 진화과정 즉, 자연선택과 유전법칙을 모방한 알고리즘이다. 유전알고리즘의 특징은 다수의 해들로 이루어진 집단에 적자생존의 이론과 자손의 형질은 부모의 유전자로부터 내려온다는 멘델의 유전 법칙을 적용하여 세대를 거듭해가며 특정 문제에 대한 최적의 해를 찾는 것이다. 유전알고리즘의 대표적인 단계로는 적합도 계산, 선택(Selection)연산, 교차(Crossover)연산, 돌연변이(Mutation)연산이 있다.

유전알고리즘을 통해 특정문제에 대한 효과적인 해를 탐색하기 위해서는 먼저 도출된 해가 얼마나 적합한지 평가하기 위한 적합도 함수가 필요하다. 적합도 계산 단계는 도출된 해가 설정된 적합도 함수에 얼마나 적합한지를 계산하는 과정으로써 이를 통해 우수한 해의 유전자는 자손세대에 전달하고 나쁜 유전자의 해는 도태시킬 수 있는 선택작업을 가능케 한다. 선택 연산은 교차연산에 쓰이는 두 개의 부모해를 선택하기 위한 연산이다. 이 선택 연산에서는 특정 문제에 대한 적합도가 높은 해가 교차연산에 선택될 확률을 높게 설정한다. 교차연산은 선택연산을 통해 선택된 두해의 특정 부분을 결합하여 하나의 새로운 해를 만들어 내는 연산이다. 교차연산을 통해 만들어진 해의 유전자는 두 부모해의 유전자로부터 물려 받은 것이다. 유전알고리즘에서는 돌연변이연산을 통해 부모해에 존재하지 않는 유전자를 자식해에 포함시켜 다양한 해를 탐색할 수 있다. 돌연변이 연산 과정에서는 돌연변이 발생 확률에 따라 해의 염색체 중 일부를 임의로 선정하여 염색체를 이루고 있는 비트 중 일부를 임의로 변형시킨다. 돌연변이연산이 일어날 확률을 높이면 보다 다양한 해를 생성할 수 있어 역동적인 최적해 탐색이 가능하지만 최적해로의 수렴속도가 떨어진다는 단점이 존재한다. 자연의 진화과정을 모방한 유전알고리즘은 쉽게 최적의 해를 찾기 어려운 시스템 최적화, 신경망 최적화, 자동제어 등의 여러 분야에 널리 적용되고 있다.

3. 연구설계

본 연구에서는 기업실적예측모델의 예측변수로 주로 활용되는 기업의 재무지표 및 실적지표들을 대상으로 주성분분석을 통해 새로운 예측변수 생성하고 그 유용성에 대해 분석하고자 한다. PCA를 통해 생성된 예측변수의 유용성을 실증적으로 분석하기 위해 기업의 재무정보를 수집하고 이를 바탕으로 유전알고리즘과 인공신경

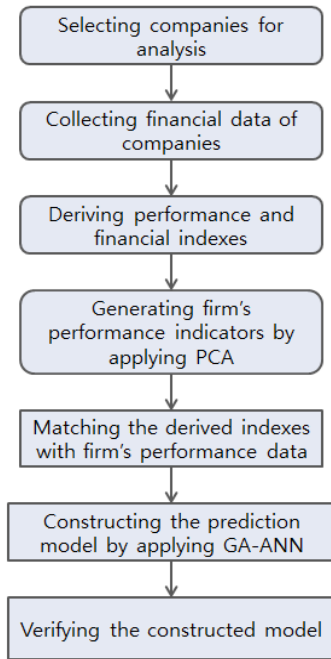


그림 1. 연구개요

Fig. 1. Summary of proposed prediction model

망을 활용한 기업실적예측 모델을 구축하고 그 성능을 분석한다. 그림.1은 본 연구의 개요이다.

먼저 본 연구에서 제안한 예측모델 구축을 위해 기업의 재무정보와 실적정보를 바탕으로 다양한 기업실적 지표를 도출한다. 그 후, 다양한 예측변수들을 예측모델 학습에 활용하게 될 경우 발생하게 되는 변수들 간의 교호작용과 상관성 등의 문제를 해결하기 위해 주성분 분석을 수행하여 새로운 예측변수를 생성한다.

1년 전 또는 직전 분기의 기업 데이터를 통해 기업실적을 예측하는 경우 현시점과 예측시점 사이의 시간차가 짧기 때문에 기업 실적의 변동 폭이 크지 않아 예측모델의 예측성능을 파악하기가 어렵다. 이 때문에 본 연구에서는 제안된 예측모델의 예측성능을 더 잘 분석하기 위해 4년 전의 기업 실적 및 재무정보를 바탕으로 기업의 실적을 예측하고자 한다. 본 연구에서는 이와 같은 시차를 적용하

기 위해 아래와 그림.2와 같은 방법으로 시차변수를 생성하여 모델을 학습시킨다.

Year	index1	index2	Year	Revenue	Profit
2011	7	8	2011	350	10
2010	4	6	2010	320	5
2009	9	9	2009	300	20
2008	8	7	2008	200	10
2007	15	18	2007	190	2
2006	5	3	2006	180	3

그림 2. 시차변수 생성의 예

Fig. 2. Example of generating time-lag effected variables

제안된 예측모델 구축을 위해 본 연구에서는 인공신경망 알고리즘을 활용하였다. 그리고 보다 나은 예측성능을 위해 유전법칙 및 자연선택과 같은 자연의 진화과정을 모방한 유전알고리즘(Genetic Algorithm)을 인공신경망 학습과정에 적용하였다. 본 연구에서 유전알고리즘은 기업실적의 실제 값과 인공 신경망을 활용하여 구축된 예측모델을 통해 도출된 기업실적 예측치 간의 MSE(Mean Squared Error)값을 최소화 하도록 신경망 노드들 간의 가중치를 조절한다.

본 연구에서 제안한 기업실적예측 모델의 성능을 검증하고 분석하기 위해 검증데이터를 활용하여 구축된 모델을 통해 도출된 기업실적의 예측치와 실제값을 비교, 분석함으로써 기업실적 예측모델의 예측성능을 검증하고 분석한다.

4. 실험 및 결과

본 연구에서 제안한 주성분 분석을 통한 기업실적 예측모델의 예측성능을 실증적으로 분석하기 위해 미국의 전기전자 기업을 대상으로 기업데이터를 수집하여 제안된 모델을 구축하였다. 모델 구축

표 1. 분석대상기업의 기술통계량(n=612, m=million)

Table 1. Descriptive statistics of the 14 financial indicators of the 51 companies

	Min	Max	Average	Standard deviation
Revenue(m)	19.07	170910.00	12812.58	23865.76
Operating profit(m)	-11057	55241.00	2001.82	5356.63
Net profit(m)	-12650	41733.00	1496.95	4218.23
Total asset(m)	5.54	207000.00	15553.62	29509.50
Total liabilities(m)	2,518	103431.00	7791.62	16918.47
Shareholders equity(m)	-1600,20	123549.00	7761.99	15505.96
Total asset growth %	-75.07	512.08	9.24	34.16
Capital gains yield %	-671.03	538.73	9.22	62.11
Growth rate of sales %	-65.70	408.54	8.50	27.59
Growth rate of net income %	-37769.57	8119.48	-83.57	1803.32
Gross-profit ratio %	-668.44	106.53	5.67	33.51
Profit rate of net worth %	-2463.78	5312.58	10.01	246.80
Debt ratio(%)	-12458.59	5940.94	106.61	649.79
Capital turnover(%)	-17664.72	4926.78	243.01	918.69

을 위해 51개의 미국 전기전자 제조기업을 대상으로 2002년부터 2013년도 사이의 재무정보와 기업실적 정보를 수집하였다.

수집된 기업의 재무정보와 실적정보를 바탕으로 총 14개의 기업 실적 예측지표를 도출하였다. 다음 표1은 모델 구축에 사용된 14개 실적 예측변수들의 기술통계를 나타낸다.

본 연구에서는 실제 기업데이터를 활용하여 제안한 예측모델을 구축하고 또 검증하기위해 수집한 데이터를 예측모델 구축을 위한 학습데이터와 예측성능 검증을 위한 검증데이터로 나누어 실험을 진행하였다. 앞서 기술한 바와 같이 본 연구에서는 특정연도의 재무지표와 기업실적 지표를 활용하여 4년 뒤의 기업 매출액을 예측하는 모델을 구축하고자 한다. 이 때문에 기업실적 예측변수와 매출액 사이에 4년의 시차를 주기위해 모델학습을 위한 학습데이터로는 2002~2008년도의 기업실적 예측변수와 2006~2012년도의 매출액이 사용되었다. 그리고 검증데이터로는 2009년도의 기업실적 예측변수와 2013년도의 매출액이 사용되었다.

표 2. 예측변수들의 공차와 VIF 값
Table 2. Tolerance and VIF values of the predictors

	Tolerance	VIF
Revenue	0.085	11.773
Operating profit	0.023	42.876
Net profit	0.025	39.906
Total liabilities	0.041	24.381
Shareholders equity	0.145	6.897
Total asset growth	0.297	3.365
Capital gains yield	0.522	1.916
Growth rate of sales	0.398	2.512
Growth rate of net income	0.958	1.044
Gross-profit ratio	0.882	1.133
Profit rate of net worth	0.165	6.072
Debt ratio	0.045	22.256
Capital turnover	0.059	16.862

매출액을 예측하기위해 도출한 14개의 예측변수들을 모델학습에 사용하기에 앞서 VIF(Variance Inflation Factor)를 통해 변수들 간의 다중공선성을 진단해보았다. 그 결과는 표.2 와 같다. VIF 값을 통해 다중공선성을 살펴본 결과 도출된 상당수의 예측변수들 사이에 다중공선성의 문제점이 있는 것으로 나타났다. 이와 같은 문제를 해결하기 위해 본 연구에서는 주성분분석을 활용하여 새로운 기업실적 예측변수를 생성하였다. 제안된 기업실적 예측모델을 구축하기 위한 주성분은 누적 주성분 설명력이 95%이상인 상위 3개의 주성분(Principal Component)을 활용하였다. 선택된 상위 3개 주성분 간의 다중공선성을 VIF값을 통해 진단해 본 결과는 다음 표.3과 같다. 표.3에서와 같이 3개의 주성분 사이에 다중공선성의 문제는 없는 것으로 나타났다.

표 3. 상위 3개 주성분의 공차와 VIF 값
Table 3. Tolerance and VIF values of top 3 PCs

	Tolerance	VIF
PC1	1,000	1,000
PC2	1,000	1,000

PC3	1,000	1,000
-----	-------	-------

본 연구에서는 상위 3개의 주성분 분석을 예측변수로 활용하여 4년 후의 기업매출액을 예측하는 예측모델을 구축하기위해 인공신경망 알고리즘을 활용하였다. 모델학습을 위해 사용한 신경망의 구조는 주성분분석을 통해 생성된 예측변수를 입력받는 3개의 입력노드, 5개의 은닉노드 그리고 매출액 예측치를 출력하는 1개의 출력노드로 이루어진 단층 신경망이다. 본 연구에서는 신경망 모델 최적화를 위해 유전알고리즘을 활용하였다. 유전알고리즘은 학습데이터를 통한 인공신경망 학습 시 신경망의 각 노드들 간의 가중치를 조절하여 예측치와 목표값 사이의 mse를 최소화 하도록 설정하였다.

이와 같이 학습된 예측모델의 예측성능을 실증적으로 검증하기 위해 본 연구에서는 2009년도의 기업실적 및 재무데이터를 활용하여 구축된 예측모델을 통해 2013년도의 기업 매출액을 예측하고 실제 기업매출액과 비교해보았다.

그 결과는 다음 그림.3과 같다. 실제 매출액과 본 연구에서 제안된 모델을 통해 예측된 매출액 예측치 사이의 MSE값은 3525.14(million\$)로 나타났다. 본 연구에서는 PCA를 통한 기업실적 예측변수생성의 효과를 비교하기 위해 PCA를 통한 입력변수들의 차원축소 과정을 제외한 유전-신경망 예측모델을 같은 방법과 같은 학습데이터를 통해 구축하고 검증데이터를 통해 2013년도의 매출액을 예측해 보았다. 그 결과 매출액 예측치와 실제 매출액 사이의 MSE는 6120,812(million\$)로 나타났다.

5. 결론 및 향후 연구

본 연구에서는 기업의 실적을 예측하기 위한 정량적 예측모델을 제안하였다. 기업실적에 영향을 미치는 요인들은 매우 다양한데, 이를 모두 예측모델의 예측변수로 활용할 경우 변수들 사이의 상관성 및 교호작용으로 인해 예측성능이 떨어질 위험이 존재한다. 이러한 문제점을 해결하기 위해 본 연구에서는 주성분 분석을 통하여 다양한 예측변수들이 가지는 정보를 최대한 확보하는 새로운 예측변수를 생성하고 이를 모델구축에 활용하는 방법을 제안하였다. 모델 학습을 위한 알고리즘으로는 인공신경망을 활용하였으며 유전알고리즘을 통해 신경망의 노드 간 가중치를 조절하였다. 제안된 실적 예측 모델을 통해 기업의 매출액을 예측한 결과 큰 어닝쇼크(Earning shock)나 어닝서프라이즈(Earning surprise) 없이 양호한 예측성능을 보였다. 그리고 PCA를 통한 예측변수들의 차원축소과정을 거치지 않은 예측모델과 본 연구에서 제안한 모델의 예측성능을 비교한 결과 PCA를 통한 변수들의 차원축소과정을 거친 예측모델의 예측성능이 더 우수한 것으로 나타났다. 이를 통해 PCA를 통한 기업실적예측 변수 생성이 다양한 지표를 사용할 시 나타나는 다중공선성의 문제를 해결하는 유용한 방법인 것을 알 수 있었다.

본 연구에서는 기업의 실적을 예측하기 위해 기업의 재무정보 및 재무적인 실적정보만을 활용하였다. 최근 들어 기업의 기술력이 기업의 실적에 큰 영향을 미치게 되었기에 향후 연구에서는 기업의 기술력을 정량적으로 나타낼 수 있는 예측변수를 도출하고 이를 예측모델 구축에 활용할 필요가 있다. 또 제안된 모델에서 사용된 인공신경망 외에 SVM과 같은 다양한 학습알고리즘을 활용하여 기업실적 예측모델을 구축하고 그 예측성능을 비교, 분석할 필요가 있을 것으로 보인다.

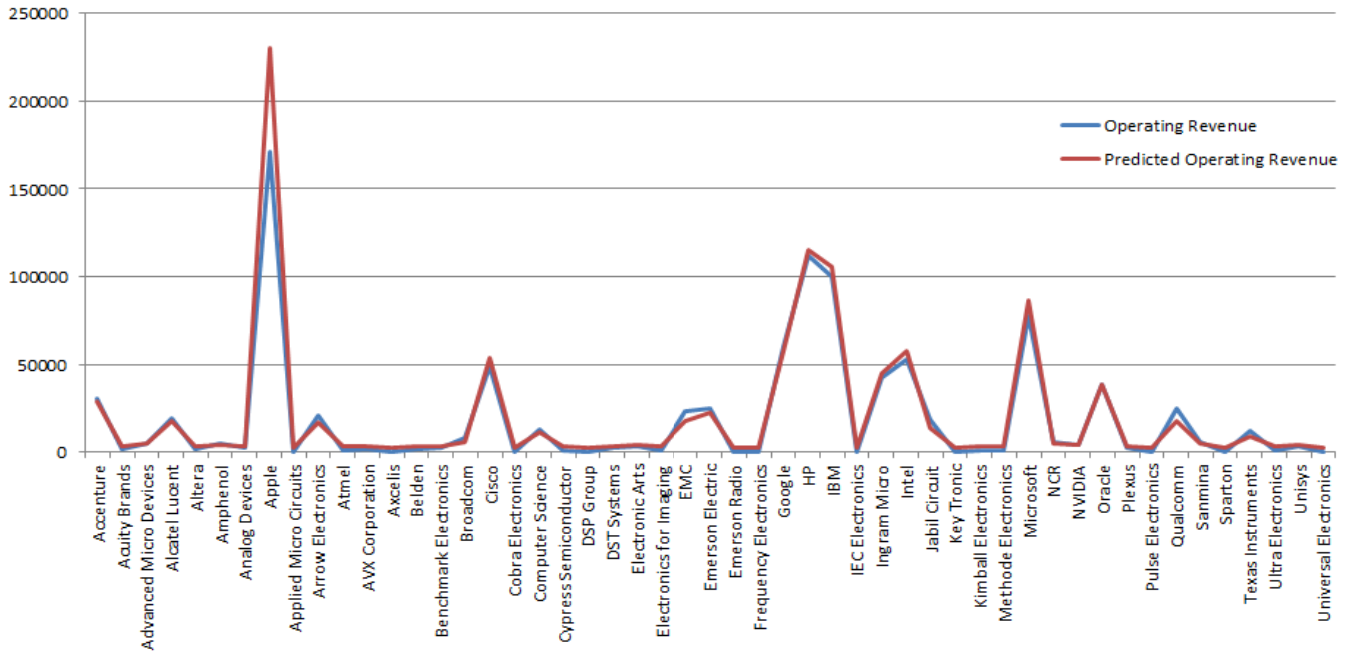


그림 3. 제안된 예측모델을 통한 매출액 예측 결과
 Fig. 3. Prediction results of the constructed prediction model(Total sales)

References

- [1] Chih-Hung Wu, Gwo-Hshiang Tzeng, Yeong-Jia Goo, Wen-Chang Fang, "A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy," *Expert Systems with Applications*, vol. 32, no. 2, pp. 397-408, Feb. 2007.
- [2] Jae H. Min, Young-Chan Lee, "Bankruptcy prediction using support vector machine with optimal choice of kernel function paramet," *Expert Systems with Applications*, vol. 28, no. 4, pp. 603-614, May. 2005.
- [3] Fengyi Lin, Ching Chiang Yeh, Meng Yuan Lee, "A Hybrid Business Failure Prediction Model Using Locally Linear Embedding And Support Vector Machines," *Romanian Journal for Economic Forecasting*, no. 1, pp. 82-97, 2013.
- [4] Yen-Yoo You, Jae-Whak Roh, "A Study on Selecting Model for Small and Medium Management Innovative Manufacturers," *Journal of Society for e-Business Studies*, vol. 15, no. 2, pp. 55-75, May. 2010.
- [5] Young Geun Shin, Sang Sung Park, Dong Sik Jang, "R&D Indicators of a Firm as Predictors for Predicting Firm Performance," *Information-An International Interdisciplinary Journal*, vol. 15, no. 2 pp. 577-596, Feb. 2012.
- [6] Joonhyuck Lee, Gabjo Kim, Dongsik Jang, Sangsung Park, "A Novel Method for Technology Forecasting Based on Patent Documents," *Advances in Intelligent Systems and Computing*, Vol. 271, pp. 81-90, 2014.
- [7] Jun, S.H., Park, S.S., Jang, D.S., "Technology forecasting using matrix map and patent clustering," *Industrial Management & Data Systems*, vol. 112, no. 5, pp. 786-807, Feb. 2012.

저 자 소개



이준혁(Joonhyuck Lee)

2012년 : 한국항공대학교 정보통신공학과 공학사
 2014년 : 고려대학교 산업경영공학부 공학석사
 2014년~현재 : 고려대학교 대학원 산업경영공학부 박사과정

관심분야 : Prediction Model, Management of Technology
 Phone : +82-2-3290-3900
 E-mail : iguana751@korea.ac.kr



김갑조(Gabjo Kim)

2011년 : 숭실대학교 산업정보시스템공학과 공학사
 2011년~현재 : 고려대학교 대학원 산업경영공학부 석·박사통합과정

관심분야 : Technology Forecasting, Data mining
 Phone : +82-2-3290-3900
 E-mail : kkjjo@korea.ac.kr



박상성(Sangsung Park)

2006년 : 고려대학교 산업시스템정보공학과
공학박사

2006년~2014년 : 고려대학교 산업경영공학부
연구교수

2014년~현재 : 한국지식재산전략원 평가위원

2014년~현재 : IP창조기업협의회 연구분과장

2014년~현재 : 고려대학교 지식재산학과 연구교수

관심분야 : Patent Analysis, Data Mining, Management of
Technology, Technology evolution

Phone : +82-2-3290-3900

E-mail : hanyul@korea.ac.kr



장동식(Dongsik Jang)

1979년 : 고려대학교 산업공학과 공학사

1985년 : 텍사스 주립대학 산업공학과
공학석사

1988년 : 텍사스 A&M 산업공학과
공학박사

1989년~현재 : 고려대학교 산업경영공학부 교수

관심분야 : PM, Pattern Recognition, Data Mining

Phone : +82-2-3290-3387

E-mail : jang@korea.ac.kr