

## 의사결정나무분석을 이용한 컨테이너 수출입 물동량 예측

### Forecasting Export & Import Container Cargoes using a Decision Tree Analysis

손용정\* · 김현덕\*\*

Yongjung Son · Hyunduk Kim

**Abstract** : The of purpose of this study is to predict export and import container volumes using a Decision Tree analysis. Factors which can influence the volume of container cargo are selected as independent variables; producer price index, consumer price index, index of export volume, index of import volume, index of industrial production, and exchange rate(won/dollar). The period of analysis is from january 2002 to December 2011 and monthly data are used. In this study, CRT(Classification and Regression Trees) algorithm is used. The main findings are summarized as followings. First, when index of export volume is larger than 152.35, monthly export volume is predicted with 858,19TEU. However, when index of export volume is between 115.90 and 152.35, monthly export volume is predicted with 716,582TEU. Second, when index of import volume is larger than 134.60, monthly import volume is predicted with 869,227TEU. However, when index of export volume is between 116.20 and 134.60, monthly import volume is predicted with 738,724TEU.

**Key Words** : Container, Export, Import, Decision Tree Analysis

---

▷ 논문접수: 2012.11.09   ▷ 심사완료: 2012.12.17   ▷ 게재확정: 2012.12.24

\* 조선대학교 무역학과 강사, 01046330773@hanmail.net, 010)4633-0773, 대표집필

\*\* 순천대학교 물류학과 교수, hdkim@sunchon.ac.kr, 010)3396-3373, 교신저자

## I. 서론

세계 컨테이너물동량은 글로벌 경기 침체로 2009년 약 4억 5,000만TEU, 2010년에는 약 5억 200만TEU를 처리하여 2009년 대비 11.5% 급증하였다. 2011년에는 전년대비 8% 증가한 약 5억 6,150만TEU를 처리한 것으로 나타나고 있다.

2012년 컨테이너 물동량의 성장률은 전년 대비 둔화될 것으로 예상되나 환적화물의 증가, 경기회복 기대감, 벌크화물의 컨테이너화 지속, 특히 신흥시장의 경제성장과 인프라 투자가 컨테이너 물동량 발생에 긍정적인 영향을 미칠 것으로 예상된다.

한편, 국내 항만의 컨테이너 화물 처리량은 2011년 약 21,610천TEU를 처리하여 전년대비 11.6% 증가하였다. 세계 경기침체로 수출입 물동량 증가세는 2010년 19.1%에서 2011년 8.6%로 둔화되었으나, 총 화물처리량은 약 21,610천TEU를 기록하여 20,000천TEU를 상회하였다. 2012년 컨테이너 물동량은 11월말 현재 약 20,152천TEU를 처리하여 2011년 대비 소폭의 증가가 예상된다.

국제적으로 정치적 불안과 사회·경제적인 혼란 상황이 일어나고 있다. 북아프리카와 중동지역의 정치 불안, 포르투갈, 아일랜드, 이탈리아, 그리스, 스페인의 재정위기는 세계 경제의 불확실성을 더욱 증폭시키고 있다. 유가 상승 역시 무역전망을 더욱 어둡게 하고 있으며, 유가상승은 기업의 투자와 가계 지출 그리고 컨테이너 물동량에 직접적인 영향을 미칠 수 있다.

세계경기 침체로 전반적으로 경제성장세가 다소 둔화되었으나, 장기적인 관점에서 세계 교역량은 상승할 것으로 전망된다. 또한 중국항만을 중심으로 한 신흥국의 경제성장 추세가 아직 이어지고 있어 환적화물을 적극 유치하고, 항만배후물류단지가 활성화되면 컨테이너 수출입 물동량도 점차 증가할 것으로 예상된다.

한편, 의사결정나무분석을 이용한 수출입 컨테이너 물동량 예측은 국내에서 다루어지지 않아 본 연구는 의사결정나무분석을 이용하여 컨테이너항만 물동량 중 수출입물동량 예측을 실시하고자 한다.

## II. 이론적 배경

### 1. 연도별 컨테이너 화물 처리 현황

연도별 컨테이너 화물 처리량 변동 추이를 살펴보면, 2011년 컨테이너 처리 물동량은 2010년에 비해 11.6% 증가한 21,611천TEU를 처리하였다. 세계 경기침체로 수출입 증

의사결정나무분석을 이용한 컨테이너 수출입물동량 예측

가세는 2010년 19.1%에서 2011년 8.6%로 둔화되었으나, 총 화물처리량은 21,610천 TEU로 2,000만TEU를 돌파하였다.

수출입 물동량은 2011년 약 13,413천TEU를 처리하여 전년대비 8.6%가 증가하였고, 환적물동량은 2010년(3,314천TEU)대비 16.2% 증가한 7,719천TEU를 기록하였다.

<표 1> 컨테이너 화물 처리현황

(단위 : 천TEU, %)

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
전 체 (증감율)	9,116 (18.6)	9,990 (9.6)	11,890 (19)	13,185 (10.9)	14,523 (10.1)	15,216 (4.8)	15,965 (4.9)	17,544 (9.9)	17,927 (2.2)	16,341 (-8.8)	19,369 (18.5)	21,610 (11.6)
수출입계	6,389	6,591	7,355	8,183	9,025	9,412	10,123	11,254	11,605	10,369	12,350	13,412
수 출 (증감율)	3,194 (10.4)	3,285 (2.8)	3,710 (12.9)	4,072 (9.8)	4,506 (10.7)	4,684 (4)	5,062 (8.1)	5,602 (10.7)	5,753 (2.7)	5,242 (-8.9)	6,191 (18.1)	6,657 (7.5)
수 입 (증감율)	3,195 (12.6)	3,306 (3.5)	3,645 (10.3)	4,111 (12.8)	4,519 (9.9)	4,728 (4.6)	5,061 (7.1)	5,652 (11.7)	5,853 (3.5)	5,128 (-12.4)	6,159 (20.1)	6,755 (9.7)
환 적 (증감율)	2,454 (47.7)	3,111 (26.8)	4,205 (35.2)	4,598 (9.3)	5,158 (12.2)	5,532 (7.2)	5,673 (2.6)	6,155 (8.5)	6,186 (0.5)	5,719 (-7.5)	6,641 (16.1)	7,719 (16.2)
연안화물	273	288	330	404	340	272	168	135	135	253	378	478

자료 : 국토해양부(Port-MIS: 항만관리정보시스템).

## 2. 항만별 컨테이너 화물 처리 현황

항만별 컨테이너 화물 처리량 변동 추이를 살펴보면, 부산항은 환적화물의 증가로 1,500만TEU를 달성하였으며, 9년 연속 세계 5위의 위상을 유지하고 있다.

수출입물동량에서 부산항은 주요 교역국가와의 수출입 물동량이 꾸준히 증가하면서 2011년 8,708천TEU를 기록하여 전년대비 11.1% 증가하였다. 광양항은 유럽 재정위기에 따른 주요 선진국과의 교역량 감소로 2010년(1,725천TEU) 대비 0.6% 감소한 1,714천TEU, 인천항은 대 중국 수출입화물 증가에 힘입어 5.1% 증가한 1,931천TEU를 기록하였다.

2011년 기준 환적 물동량을 보면, 부산항은 북중국항만 대체 환적이 꾸준히 증가하면서 전년대비 17.1% 증가한 7,353천TEU, 광양항은 전년대비 4.5%(328천TEU)증가하였고, 인천항은 전년대비 12.0%(18천TEU) 감소한 것으로 나타나고 있다.

<표 2> 항만별 컨테이너 화물 처리실적

(단위 : 천TEU)

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
부산항	7,540	8,073	9,453	10,408	11,492	11,843	12,039	13,261	13,453	11,980	14,194	16,185
수출입	5,035	5,011	5,522	6,035	6,595	6,579	6,803	7,444	7,638	6,569	7,836	8,708
환적	2,390	2,943	3,887	4,251	4,792	5,179	5,208	5,811	5,808	5,372	6,276	7,353
연안화물	116	119	44	122	105	85	28	7	7	39	82	124
광양항	642	855	1,080	1,185	1,322	1,441	1,756	1,723	1,810	1,810	2,073	2,073
수출입	551	645	718	803	956	1,097	1,306	1,406	1,488	1,478	1,725	1,714
환적	64	166	314	344	360	343	448	314	322	306	314	328
연안화물	27	44	48	38	6	1	2	3	0	27	34	31
인천항	611	663	770	821	935	1,149	1,377	1,664	1,703	1,578	1,903	1,998
수출입	483	537	651	710	828	1,056	1,331	1,628	1,668	1,523	1,838	1,931
환적	-	1	2	3	5	6	13	18	24	19	21	18
연안화물	128	125	117	108	102	86	33	18	12	36	44	48
울산항	236	258	277	318	303	316	337	380	401	319	336	326
수출입	236	258	277	318	301	313	333	371	373	302	327	322
환적	-	-	-	-	2	4	4	10	28	16	3	2
평택,당진항	1	21	66	152	190	227	260	319	356	378	447	530
수출입	1	21	66	152	190	227	259	317	353	371	426	502
환적	-	-	-	-	-	-	1	2	3	5	9	9

자료 : 국토해양부 (Port-MIS: 항만관리정보시스템).

### 3. 선행연구

데이터 마이닝은 일반적으로 데이터베이스에 있는 방대한 양의 정보에 숨어 있는 지식을 자동적으로 추출하는 과정이다. 데이터 마이닝은 방대한 데이터베이스로 숨어 있는 예측정보를 추출하는 것이다(Roiger, R. and M. Heatz, 2003). 이런 데이터 마이닝 기법에는 기술모형과 예측모형으로 분류할 수 있다. 기술모형에 패턴분석, 군집분석과 연관성 분석이 있고, 예측모형에 회귀분석모형, 의사결정나무분석, 인공신경망분석, 사례기반추론, 유전자 알고리즘, 퍼지이론 등이 있다.

예측을 위한 통계기법 중 가장 대표적인 것은 회귀모형과 시계열모형이 있으며, 그

외에도 로지스틱회귀모형, 판별분석, 일반화 선형모형 등 많은 기법이 있다. 시계열모형을 이용한 물동량 예측에 관한 연구는 국내에서 다양하게 이루어져 왔으며, 데이터마ining 기법 중에서 인공신경망분석을 이용한 전찬영·송주미(2007)의 연구가 있다.

따라서 본 연구는 국내에 물동량 예측에 다루어지지 않은 의사결정나무분석을 이용하여 항만물동량 예측을 실시해보자 한다.

전찬영·송주미(2007)는 수출입품목 각각 10개를 선정하여, 4개의 변동유형으로 군하여 회귀분석과 인공신경망모형의 우수성을 적합성, 신뢰성, 강건성 측면에서 각각 비교·분석하였다. 분석결과 물동량이 전반적으로 지속적인 상승을 하는 ‘상승지속형’의 경우에는 적합성, 신뢰성, 강건성 모든 측면에서 회귀분석이 인공신경망보다 우수한 것으로 나타났다. 전반적으로 상승하나 연도별로 증가율이 차츰 낮아지는 ‘상승둔화형’과 정체하거나 전반적으로 증가나 감소를 하지 않는 ‘정체형 또는 불규칙형’의 경우에는 3가지측면에서 모두 인공신경망모형이 더 나은 것으로 나타났다. 구간별로 변동이 불규칙이나 단계별 저점이 높아지는 ‘불규칙 상승형’의 경우에는 적합성 측면에서 회귀분석이 신뢰성과 강건성 측면에서는 인공신경망모형이 더 좋은 것으로 나타났다.

의사결정나무분석을 이용한 연구는 이재식·이진천(2006)의 자동차보험 고객이탈 예측, 이준용·최미화·이상엽(2007)의 아파트 가격 예측, 조준희·강부식(2007)의 코스닥 기업 도산예측, 이용구·허준(2009)의 생명보험사의 개인연금 보험예측, 장철호(2009)의 공정무역상품 소비활성화 전략수립 연구, 정지혜·김태훈·김상열(2010)의 중소기업 도산예측모형, 최상일·박태진·강정규(2011)의 주거유형선택 예측 등 다양하게 연구되고 있다.

### Ⅲ. 연구방법 및 결과

#### 1. 의사결정나무분석

의사결정나무모형은 데이터베이스에 존재하는 방대한 양의 자료로부터 사전에 알려지지 않은 암시적이고 유용한 정보를 추출하는 것으로 인공지능 분야의 기계학습 이론에 근거한다. 의사결정나무모형은 분석대상을 몇 개의 소집단으로 분류하는 규칙을 나무구조로 만들어 이를 새로운 대상에 적용함으로써 그 특성을 예측하는 모형으로 특히 해석이 용이하고 분류와 예측이 쉬워 매우 유용한 도구로 사용되고 있다. 의사결정나무모형에 의한 결과 규칙은 바로 우리가 쓰는 말로 표현되어 쉽게 이해할 수 있다(장철호, 2009).

<표 3> 의사결정나무분석의 용도

용도	내용
세분화	데이터를 비슷한 특성을 갖는 몇 개의 그룹으로 분할하여 각 그룹별 특성을 발견하는 경우 또는 각 고객이 어떤 집단에 속하는지를 파악하고자 할 경우
분류	관측개체를 여러 예측변수들에 근거하여 목표변수의 범주를 등급으로 분류하고자 하는 경우
예측	자료로부터 규칙을 찾아내고 이를 이용하여 미래의 사건을 예측하고자 하는 경우
변수선택	매우 많은 수의 예측변수 중에서 목표변수에 큰 영향을 미치는 변수들을 골라내고자 하는 경우
교호작용효과	여러개의 예측변수들이 결합하여 목표변수에 작용하는 규칙을 파악하고자 할 경우
범주병합 또는 연속변수의 범주화	범주형 목표변수의 범주를 소수의 몇 개로 병합하거나 연속형 목표변수를 몇 개의 등급으로 이산화하고자 하는 경우

자료 : 최중후 · 한상태 · 강현철 · 김은석, 『Answer Tree를 이용한 데이터마이닝 의사결정나무분석』, SPSS아카데미, 1998, p.18.

의사결정나무모형은 의사결정규칙을 도표화하여 관심대상이 되는 집단을 몇 개의 소집단으로 분류하거나 예측하는 분석방법이다. 분석과정이 나무구조에 의해서 표현되기 때문에 분류 또는 예측을 목적으로 하는 다른 방법들에 비해 연구자가 분석과정을 쉽게 설명할 수 있는 장점이 있다.

1980년대 이후 의사결정나무분석에 대한 알고리즘은 계속적으로 개발되고 있으며, 분석의 기본 목적과 자료의 구조에 의해서 분석방법이 구분된다.

분석방법에는 CHAID(Chi-squared automatic interaction detection; Kass, 1980), Exhaustive CHAID(Biggs et al., 1991), CRT(Classification and Regression Trees; Breiman et al., 1984), QUEST (Quick, Unbiased, Efficient Statistical Tree; Loh and shin, 1997)등이 있다.

<표 4> 의사결정나무분석의 종류

	CHAID	Exhaustive CHAID	CRT	QUEST
목표변수	질적변수, 양적변수	질적변수, 양적변수	질적변수, 양적변수	명목형 질적변수
예측변수	질적변수, 양적변수	질적변수	질적변수, 양적변수	질적변수, 양적변수
분리기준	F검정, 카이제곱통계량	F검정, 카이제곱통계량	지니계수감소	F검정, 카이제곱통계량
분리개수	다지분리	다지분리	이지분리	이지분리

자료 : 강병서 · 김계수, 『SPSS 17.0 사회과학통계분석』, 한나래아카데미, 2009, p.545.

## 2. 자료수집 및 변수선정

<표 1>과 <표 2>에서 제시한 바와 같이, 컨테이너 물동량은 수출, 수입, 환적, 연안 화물 등으로 구성되는데, 본 연구는 의사결정나무분석을 이용하여 컨테이너 수출입 물동량을 예측하고자 한다.

변수선정은 국내 수출입 물동량 예측과 관련된 문헌 중 모수원(2001), 장문철(2010)의 연구를 참고하여 컨테이너 수출입 물동량의 변화에 영향을 미칠 가능성이 있는 변수를 선정하였다.

물가를 반영하는 변수로 생산자물가지수와 소비자물가지수를 선정하였는데, 생산자물가지수는 국내생산자가 공급하는 모든 상품 및 일부 서비스의 가격수준 변동을 측정하는 지표로서 상품 및 서비스의 수급상황 파악, 경기동향 판단지표 및 GDP 디플레이터 등으로 이용되고, 소비자물가지수는 소비생활을 위하여 구입하는 각종 상품과 서비스의 가격 변동을 파악하기 위하여 작성되는 지표로서 소비자가 일정한 생활수준을 유지하는 데 필요한 소득내지 소비금액의 변동을 나타내기 때문에 소비자의 구매력과 생계비 등의 확인할 수 있는 지표이다(한국은행 웹사이트(www.bok.or.kr)).

수출입 물량지수는 무역지수 중 수출입물량의 변동을 나타내 주는 것이고, 세계경기와 국내경기를 반영하기 위해 대리변수로 미국과 한국의 산업생산지수를 이용하였는데, 산업생산지수는 기업의 연간 생산량, 출하량 및 재고 변동량을 나타내는 지수로 업종별, 기업별 경기를 측정하고 예측하는 경제 지표이다. 그리고 원/달러 환율은 한국수출상품의 가격경쟁력에 영향을 미칠 수 있기 때문에 선정하였다.

분석기간은 2002년 1월부터 2011년 12월까지 10년간의 월별자료를 이용하였으며, 분석패키지는 SPSS 20.0을 이용하였다.

<표 5> 변수선정

	변수명	자료출처
종속변수	수출 컨테이너 물동량(TEU)	관세청 수출입무역통계(수출입화물통계)
	수입 컨테이너 물동량(TEU)	
독립변수	생산자·소비자물가지수(2005=100), 수출·수입물량지수(2005=100), 환율(원) 미국·한국산업생산지수(2005=100)	한국은행 경제통계시스템(ECOS)

주 : ( )는 단위

의사결정나무분석을 수행하기 위한 다양한 분리기준, 정지규칙, 가지치기 방법들이 제안되고 있으며, 이들을 어떻게 결합하느냐에 따라 서로 다른 의사결정나무 형성방법이 만들어진다. 의사결정나무를 형성하기 위해 다양한 알고리즘이 제안되고 있는데, SPSS 프로그램에서 제공되는 의사결정분석 알고리즘은 CHAID, Exhaustive CHAID, CRT, QUEST 방법 등이 있다.

본 연구는 CRT(Classification and Regression Trees; Breiman · Friedman · Olshen · Stone, 1984) 알고리즘을 활용하는 의사결정나무분석을 실시한다.

CRT방법은 종속변수에 대해서 가능한 한 많은 동질적인 그룹이 속하도록 노드를 생성하는 방법을 사용한다. 종속변수에 대해서 똑같은 값을 갖는 케이스로만 이루어진 최종노드를 순수노드라고 한다. Breiman · Friedman · Olshen · Stone(1984)는 CRT를 통하여 의사결정나무 규칙에 의한 나무모형의 성장, 최종모형의 선정을 위한 나무모형의 가지치기(pruning) 등을 이론적으로 정립하였다. CRT는 좌우 2개 가지로만 분할되는 이진분할 방법을 이용한다. 각 노드에서의 분할 규칙은 데이터가 만들어 낼 수 있는 모든 가능한 분할 규칙 중 불순도를 최소로 하는 것을 선택한다. 불순도 함수로는 지니지수, 엔트로피지수, 편차지수 등이 있다. 독립변수에 결측치가 존재하는 경우에 이 변수와 밀접한 관련성을 가지는 다른 독립변수를 이용하여 분류 작업에 이용한다(박성현 · 김성수 · 황현식, 2011).

의사결정나무분석을 위한 분석설정은 다음과 같은 정지규칙을 설정하였다. 전체적 분리정지(Maximum tree depth)는 5, 부모마디의 관측 수 제한(Minimum Number of Case for parent)은 25, 자식마디의 관측 수 제한(Minimum Number of Case for child)은 1로 설정하였다.

### 3. 분석결과

#### 1) 컨테이너 수출물동량 예측

컨테이너 수출물동량에 대한 최적분리는 수출물량지수에 의해 분리되었다. 수출물량지수는 115.90으로 분리되어 있는데, 수출물량지수가 115.90보다 큰 경우는 다시 수출물량지수가 152.35보다 큰 경우와 115.90과 152.35사이인 경우로 분리되어진다. 여기서 수출물량지수가 152.35보다 큰 경우는 858,191TEU/(월)이고, 115.90과 152.35사이인 경우는 716,582TEU/(월)로 컨테이너 수출물동량이 예측된다.

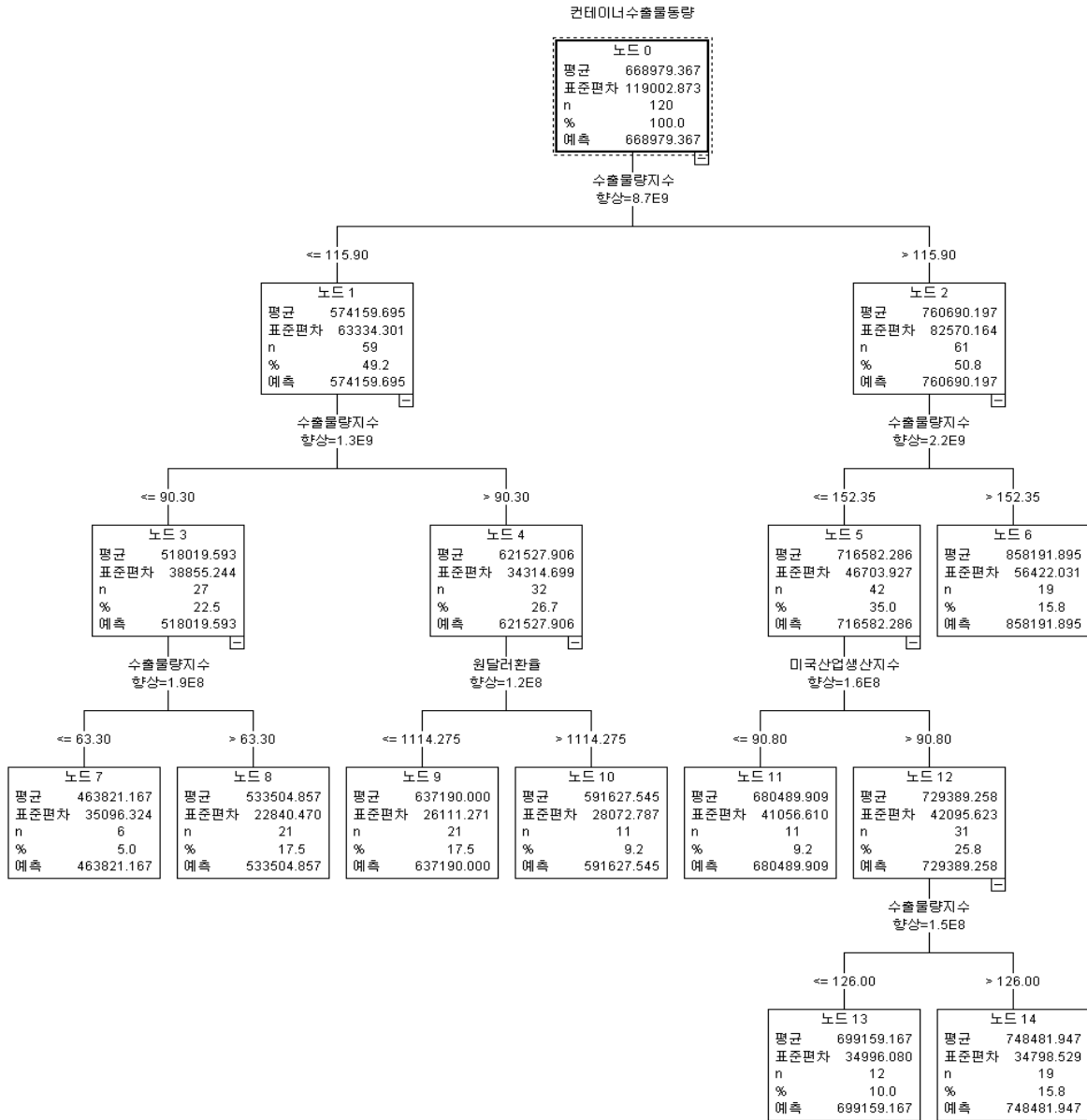
수출물량지수가 115.90보다 작은 경우는 다시 90.30보다 작은 경우와 90.30과 115.90사이인 경우로 나누어진다. 수출물량지수가 90.30과 115.90사이인 경우는 원달러환율



의사결정나무분석을 이용한 컨테이너 수출입물동량 예측

1,114원으로 분리된다. 1,114원이하인 경우는 637,190TEU/(월), 1,114월 이상인 경우는 591627TEU/(월)로 예측된다.

<그림 1> 컨테이너 수출물동량 예측



이익도표를 통해 나무구조분석에 대한 정보를 알 수 있다. 이익도표는 해당마디에서 자료가 얼마나 잘 분류되었는지를 알 수 있게 해준다. 다음의 <표 6>은 나무분리구조에 대한 이익도표의 통계량이 제시되어 있다.

<표 6> 노드에 대한 이익 요약

Node	N	퍼센트	평균
6	19	15.8	858,191.89
14	19	15.8	748,481.95
13	12	10.0	699,159.17
11	11	9.2	680,489.91
9	21	17.5	637,190.00
10	11	9.2	591,627.55
8	21	17.5	533,504.86
7	6	5.0	463,821.17

## 2) 컨테이너 수입물동량 예측

컨테이너 수출물동량과 마찬가지로 컨테이너 수입물동량에 대한 최적분리도 수입물량지수에 의해 분리되었다.

수입물량지수가 116.20에서 분리되어 지는데, 수입물량지수가 116.20보다 큰 경우는 다시 수입물량지수가 134.60보다 큰 경우와 116.20과 134.60사이인 경우로 분리되어진다. 여기서 수입물량지수가 134.60보다 큰 경우는 869,227TEU/(월)이고, 116.20과 134.60사이인 경우는 738,724TEU/(월)로 컨테이너 수입물동량이 예측된다.

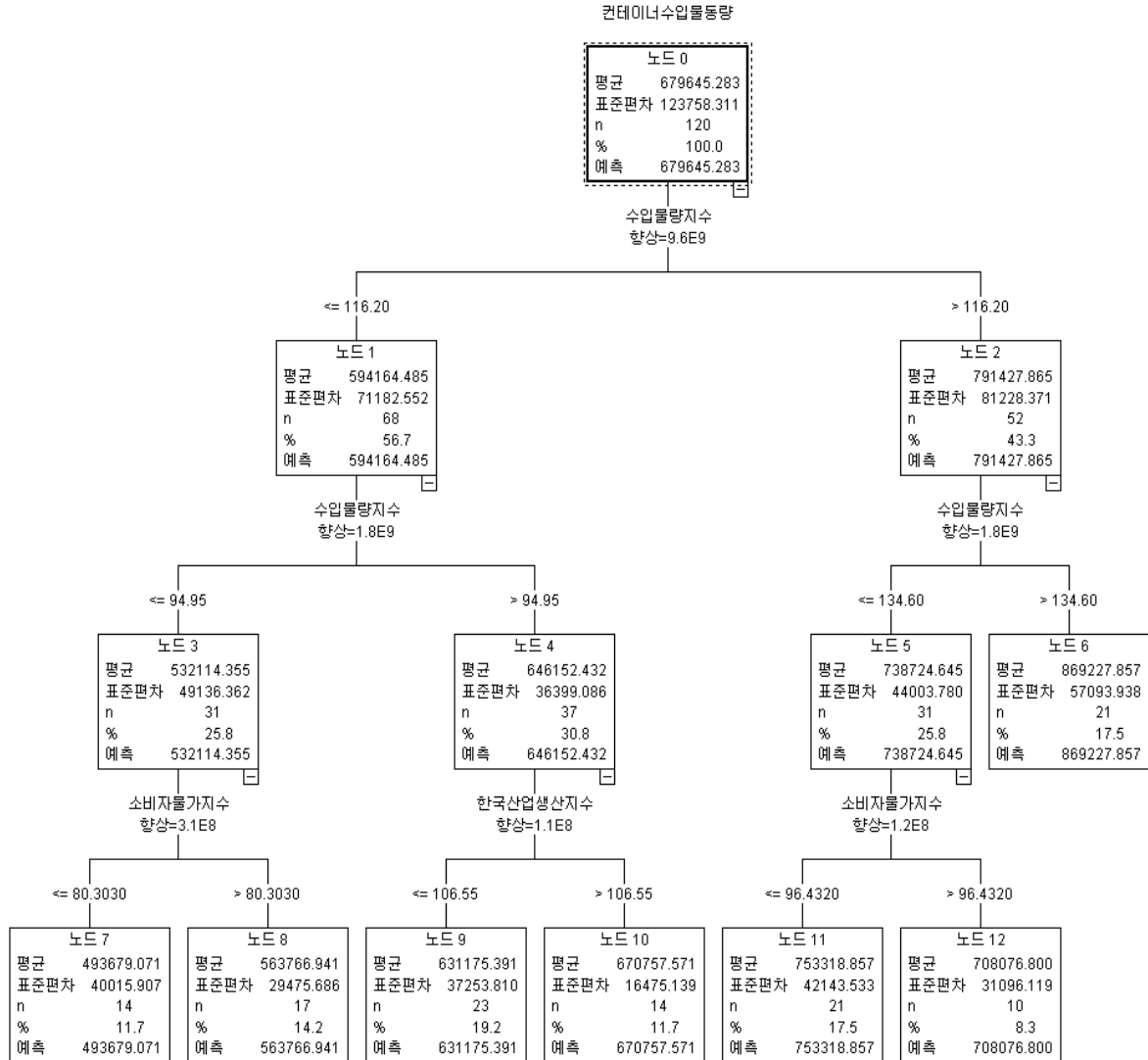
수입물량지수가 116.20보다 작은 경우는 다시 94.95보다 작은 경우와 94.95와 116.30사이인 경우로 나누어진다. 수출물량지수가 94.95와 116.30사이인 경우는 한국 산업생산지수로 분리되고, 94.95이하인 경우는 소비자물가지수로 분리되어진다.

다음의 <표 7>은 나무분리구조에 대한 이익도표의 통계량이 제시되어 있다.

<표 7> 노드에 대한 이익 요약

Node	N	퍼센트	평균
6	21	17.5	869,227.86
12	21	17.5	753,318.86
11	10	8.3	708,076.80
10	14	11.7	670,757.57
9	23	19.2	631,175.39
8	17	14.2	563,766.94
7	14	11.7	493,679.07

<그림 2> 컨테이너 수입물동량 예측



## V. 요약 및 결론

중국, 일본을 비롯한 세계 각국은 자국 항만을 허브항만으로 육성하기 위해 다각도의 노력을 기울이고 있다. 따라서 부산항과 광양항 등 국내 항만이 동북아 허브항만으로 성장하기 위해서는 환적화물 중심이 아닌 고부가가치를 창출할 수 있는 노력을 경주하고 세계 주요 항만들과 네트워크를 구축하는 노력을 계속하여야 할 것이다.

분석결과는 컨테이너 수출물동량에 대한 최적분리는 수출물량지수에 의해 분리되었다. 수출물량지수는 115.90으로 분리되어 있는데, 수출물량지수가 115.90보다 큰 경우는 다시 수출물량지수가 152.35보다 큰 경우와 115.90과 152.35사이인 경우로 분리되어진다. 여기서 수출물량지수가 152.35보다 큰 경우는 858,191TEU/(월)이고, 115.90과 152.35사이인 경우는 716,582TEU/(월)로 컨테이너 수출물동량이 예측된다.

수출물량지수가 115.90보다 작은 경우는 다시 90.30보다 작은 경우와 90.30과 115.90사이인 경우로 나누어진다. 수출물량지수가 90.30과 115.90사이인 경우는 원달러환율 1,114원으로 분리된다. 1,114원이하인 경우는 637,190TEU/(월), 1,114원 이상인 경우는 591627TEU/(월)로 예측된다.

컨테이너 수입물동량에 대한 최적분리는 수입물량지수에 의해 분리되었다. 수입물량지수가 116.20에서 분리되어 있는데, 수입물량지수가 116.20보다 큰 경우는 다시 수입물량지수가 134.60보다 큰 경우와 116.20과 134.60사이인 경우로 분리되어진다. 여기서 수입물량지수가 134.60보다 큰 경우는 869,227TEU/(월)이고, 116.20과 134.60사이인 경우는 738,724TEU/(월)로 컨테이너 수입물동량이 예측된다.

수입물량지수가 116.20보다 작은 경우는 다시 94.95보다 작은 경우와 94.95와 116.30사이인 경우로 나누어진다. 수출물량지수가 94.95와 116.30사이인 경우는 한국 산업생산지수로 분리되고, 94.95이하인 경우는 소비자물가지수로 분리되어진다.

우리나라의 항만물동량은 전통적으로 수출입 물동량이 주도하였으며, 본 연구의 결과도 수출입물량의 변동을 나타내 주는 수출입 물량지수가 수출입 물동량에 영향을 주는 것으로 분석되었다. 그러나 대외적으로 글로벌 금융위기에 따른 실물경기의 위축과 환적화물의 증가, 대내적으로 제조업의 성장세 둔화, 고령화에 의한 국내소비의 침체, 서비스 산업의 비중 확대 등 산업구조의 변화에 따라 수출입 물동량의 지속적인 감소가 예상된다.

따라서 항만당국과 항만물류기업은 연계수송망 확충 및 항만배후물류단지를 개발하여 새로운 물동량 창출을 위한 마케팅 활동이 필요하며, 이와 함께 기존고객의 이탈을 막고 신규고객을 지속적으로 유치할 수 있는 항만서비스 강화에 많은 노력이 요구된다. 중국지역 항만의 성장에 따라 항만시설의 우위와 항만비용 절감과 같은 인센티브 제공으로 인한 물동량 확보는 한계가 있으므로 항만관련 기관의 협력을 바탕으로 한 적극적인 항만세일즈 활동이 지속적인 항만물동량 확보에 필요하다.

컨테이너 물동량 중 환적화물과 연안화물은 제외하고 수출입물동량에 대해서만 예측하였고, 다양한 변수를 적용하지 못하였다. 또한 데이터 마이닝 분석 중 회귀분석모형, 인공신경망분석, 사례기반추론, 유전자 알고리즘, 퍼지이론 등이 있으나, 의사결정나무 모형만 분석한 점 등이 본 연구의 한계점이다. 추후의 연구에서는 다양한 변수의 적용, 다양한 데이터 마이닝 분석과의 비교 연구 등이 진행되어야 할 것으로 보인다.

## 참고문헌

- 강병서·김계수, 『SPSS 17.0 사회과학통계분석』, 한나래아카데미, 2009.
- 모수원, “항만의 하역능력증대를 위한 수출입 물동량 예측”, 『무역학회지』, 제26권 제1호, 2001, 275-297.
- 박성현·김성수·황현식, 『고급 SPSS 이해와 활용』, 한나래아카데미, 2011.
- 박준호·김태호·주용진, “의사결정나무분석을 이용한 벤처기업의 공간적 입지특성에 관한 연구-강남구, 금천구를 중심으로-”, 『국토지리학회지』, 제45권 1호, 2011, 175-188.
- 손용정, “항만경쟁력 제고를 위한 항만교역량 예측”, 『한국항만경제학회지』, 제25집 제1호, 2009, 1-14.
- \_\_\_\_\_, “세계 주요 컨테이너항만의 효율성 비교 연구”, 『한국항만경제학회지』, 제26집 제1호, 2010, 131-143.
- \_\_\_\_\_, “SWOT/AHP 분석을 이용한 광양항의 발전 전략에 관한 연구”, 『한국항만경제학회지』, 제27집 제1호, 2011, 247-262.
- 이용구·허준, “생명보험사의 개인연금 보험예측 사례를 통해서 본 의사결정나무 분석의 설명변수 축소에 관한 비교 연구”, 『한국데이터정보과학회지』, 제20권 제1호, 2009, 179-190.
- 이재식·이진천, “다중모델을 이용한 자동차 보험 고객의 이탈예측”, 『한국지능정보시스템학회논문지』, 제12권 제2호, 2006, 167-183.
- 이준용·최미화·이상엽, “데이터 마이닝 적용을 통한 아파트 가격 예측에 관한 연구”, 『국토계획』, 제42권 제4호, 2007, 135-148.
- 장문철, “우리나라 수출입 물동량의 구조분석과 향후 5년의 시계열 예측에 관한 연구”, 『유통정보학회지』, 제13권 제4호, 2010, 177-199.
- 장철호, “의사결정나무모형을 이용한 공정무역상품의 소비활성화 전략”, 『식품유통연구』, 제26권 제1호, 2009, 51-71.
- 전찬영·송주미, “인공신경망모형의 항만물동량 예측 적용에 관한 연구”, 『해운물류연구』, 제53호, 2007, 65-82.
- 정지혜·김태훈·김상열, “중소기업의 도산예측모형에 관한 비교연구: 기술신용보증기금의 보증지원을 받은 기업을 중심으로”, 『e-비즈니스연구』, 제11권 제5호, 2010, 211-240.
- 조경호, “공무원의 삶의 질 요인에 대한 의사결정나무분석”, 『한국인사행정학회보』, 제2권 제2호, 2003, 131-144.
- 조성빈, “의사결정나무추론을 이용한 채용결정요인 분석”, 『서강경영논총』, 제21-2집, 2010, 1-14.
- 조준희·강부식, “코스닥기업의 도산예측모형에 관한 연구”, 『산업경제연구』, 제20권 제1

- 호, 2007, 141-160.
- 최상일 · 박태진 · 강정규, “의사결정나무모형을 이용한 노년기 주거유형선택 예측에 관한 연구”, 『주거환경』, 제9권 2호(통권 제16호), 2011, 235-251.
- 최종후 · 한상태 · 강현철 · 김은석, 『Answer Tree를 이용한 데이터마이닝 의사결정나무분석』, SPSS아카데미, 1998.
- Bigg, D., B. de Ville and E. Suen, “A Method of choosing multiway partitions for classification and decision trees,” *Journal of Applied Statistics*, 18, 1991, 49-62.
- Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone, *Classification and Regression Trees*, Wadsworth, Belmont, 1984.
- IBM, *IBM SPSS Decision Trees 20*.
- Kass, G., “An exploratory technique for investigating large quantities of categorical data,” *Applied Statistics*, 29(2), 1980, 119-127.
- Loh, W. and Y. Shin. *Forthcoming. Split selection Methods for classification trees*, Statistica Sinica, Taiwan, 1997.
- Roiger, R., M. Heatz, *Data Mining: A Tutorial Based Primer*, Addison Wesley, 2003.
- [www.customs.go.kr](http://www.customs.go.kr).
- [www.ecos.bok.or.kr](http://www.ecos.bok.or.kr)
- [www.index.go.kr](http://www.index.go.kr).
- <https://portmis.go.kr>.

## 국문요약

# 의사결정나무분석을 이용한 컨테이너 수출입 물동량 예측

손용정 · 김현덕

본 연구는 의사결정나무분석을 이용하여 컨테이너 수출입 물동량을 예측한다. 컨테이너 수출입 물동량에 영향을 미칠 가능성이 있는 요인을 독립변수로 선정하였는데, 생산자물가지수와 소비자물가지수, 수출물량지수와 수입물량지수, 미국과 한국의 산업생산지수, 그리고 원/달러 환율을 선정하였다. 분석기간은 2002년 1월부터 2011년 12월까지 10년간의 월별자료를 이용하였으며, 의사결정나무를 형성하기 위해 다양한 알고리즘이 제안되고 있는데, CRT(Classification and Regression Trees)알고리즘을 활용하였다. 분석결과는 첫째, 컨테이너 수출물동량에 대한 최적분리는 수출물량지수에 의해 분리되었다. 수출물량지수는 115.90으로 분리되어 있는데, 수출물량지수가 115.90보다 큰 경우는 다시 수출물량지수가 152.35보다 큰 경우와 115.90과 152.35사이인 경우로 분리되어진다. 여기서 수출물량지수가 152.35보다 큰 경우는 858,191TEU/(월)이고, 115.90과 152.35사이인 경우는 716,582TEU/(월)로 컨테이너 수출물동량이 예측된다. 둘째, 컨테이너 수입물동량에 대한 최적분리는 수입물량지수에 의해 분리되었다. 수입물량지수가 116.20에서 분리되어 있는데, 수입물량지수가 116.20보다 큰 경우는 다시 수입물량지수가 134.60보다 큰 경우와 116.20과 134.60사이인 경우로 분리되어진다. 여기서 수입물량지수가 134.60보다 큰 경우는 869,227TEU/(월)이고, 116.20과 134.60사이인 경우는 738,724TEU/(월)로 컨테이너 수입물동량이 예측된다.

**핵심 주제어:** 컨테이너, 수출, 수입, 의사결정나무분석