http://www.jsebs.org
ISSN: 2288-3908

The Journal of Society for e-Business Studies
Vol.22, No.4, November 2017, pp.71-101
https://doi.org/10.7838/jsebs.2017.22.4.071

미국 무역정책 변화가 국내 중공업 기업의 경영성과에 미치는 영향

Predicting Performance of Heavy Industry Firms in Korea with U.S. Trade Policy Data

박진수(Jinsoo Park)*, 김경호(Kyoungho Kim)**,
김범수(Buomsoo Kim)***, 서지혜(Jihae Suh)****

초 록

미국 무역위원회(United States International Trade Commission)는 불공정 무역으로 인해 무역 질서를 해치는 경우 상계 관세(Countervailing Duties)와 반덤핑 관세(Antidumping Duties) 등을 징수하고 있다. 본 연구에서는 상기 연구 목적을 달성하기 위하여 상계 관세 및 반덤핑 관세와 관련된 데이터를 수집해 양적 분석을 수행하였다. 몇 가지 데이터 마이닝(Data mining) 기법을 활용한 본 연구의 양적 분석 결과, 미국의 상계 관세 및 반덤핑 관세 부과 경향이 우리나라의 중공업 산업의 성장률에 유의한 영향을 미친다고 잠정적으로 결론 내릴 수 있었다. 본 연구의 가장 큰 기여점은 '*미국의 보호주의 무역기조가 울산지역의 주력산업의 경영성과에 부정적인 영향을 미칠 수 있다*'는 직관적인 명제를 과거 데이터를 가지고 객관적으로 검증해보고 그 영향 정도를 계량화해 측정할 수 있도록 한 것이라고 할 수 있다.

ABSTRACT

Since late 2016, protectionism has been a major trend in world trade with the Great Britain exiting the European Union and the United States electing Donald Trump as the 45th president. Consequently, there has been a huge public outcry regarding the negative prospects of heavy industry firms in Korea, which are highly dependent upon international trade with Western countries including the United States. In light of such trend and concerns, we have tried to predict business performance of heavy industry firms in Korea with data regarding trade policy of the United States.

United States International Trade Commission (USITC) levies countervailing duties and anti-dumping duties to firms that violate its fair-trade regulations. In this study, we have performed data analysis with past records of countervailing duties and anti-dumping duties. With results from clustering analysis, it could be concluded that trade policy trends of the Unites States significantly affects the business performance of heavy industry firms in Korea. Furthermore, we have attempted to quantify such effects by employing long short-

본 논문은 한국은행의 재정지원을 받아 한국은행 울산본부와 공동으로 작성되었다.
* First/Corresponding Author, Graduate School of Business, Seoul National University(jinsoo@snu.ac.kr)
** Bank of Korea(kimkyoungho@bok.co.kr)
*** Graduate School of Business, Seoul National University(buomsookim0927@gmail.com)
**** SNU Big Data Institute(jihaesuh77@gmail.com)
Received: 2017-10-16, Review completed: 2017-11-15, Accepted: 2017-11-16

term memory (LSTM), a popular neural networks model that is well-suited to deal with sequential data. Our major contribution is that we have succeeded in empirically validating the intuitive argument and also predicting the future trend with rigorous data mining techniques. With some improvements, our results are expected to be highly relevant to designing regulations regarding heavy industry in Korea.

키워드 : 데이터 마이닝, 경영성과, 무역정책, 기계 학습, 딥 러닝, 장단기 기억 네트워크
Data Mining, Business Impact, Trade Policy, Machine Learning, Deep Learning, Long Short-Term Memory

1. 서 론

최근 들어 영국이 유럽연합에서 탈퇴하는 브렉시트(Brexit)와 미국 트럼프 대통령 취임과 함께 세계적으로 서구 경제권의 보호주의 무역 기조에 대한 우려가 강화되고 있다. 특히, 울산지역의 주력산업은 자동차, 조선 등 중화학 공업 중심의 수출 산업으로 간접 수출 효과까지 포함해 미국과의 교역에 높은 의존도를 보이고 있어 미국의 보호주의 무역 기조의 강화 추세가 울산지역의 주력산업에 악영향을 미칠 것이라는 우려의 목소리가 커지고 있다. 실제로 지난 3월 미국 정부는 현대중공업의 수출 변압기에 61%의 높은 반덤핑 관세를 부과하며 세간의 화제가 되었다[2].

미국 무역위원회(United States International Trade Commission)는 불공정 무역으로 인해 국제무역 질서를 해치고 미국경제에 악영향을 미칠 것으로 판단되는 경우 상계 관세(Countervailing Duties)와 반덤핑 관세(Anti-dumping Duties) 등을 징수하고 있다. 본 연구에서는 상계 관세와 반덤핑 관세 부과 기록을 중심으로 미국의 보호주의 무역정책이 울산지역 주력산업의 경영성과에 미치는 영향을 살펴보았다. 미국은 울산지역의 수출 상대국 중에서 큰 비중을 차지하고 있는 국가로, 미국의 보호주의적 무역정책이 울산지역 기반 산업에 부정적인 영향을 미칠 것이라는 추측은 누구나 할 수 있다. 하지만, 본 연구는 이를 과거 데이터를 가지고 객관적인 방식으로 검증해 보았고, 나아가 그러한 영향의 정도를 수치화해 측정하고 예측할 수 있도록 하였다는 점에서 중요성을 갖는다고 할 수 있다.

본 연구는 다음과 같이 이루어져 있다. 제 2장에서는 울산지역 주력산업의 전반적인 해외 수출 현황과 최근 추이에 대해서 알아본다. 구체적으로는 중화학 공업 중심의 울산지역 산업의 해외 수출액을 연도별, 국가별, 품목별로 비교 분석해 보았다. 제 3장에서는 미국 무역정책의 역사와 변화 과정을 살펴본다. 미국 건국 초기부터 보호주의 정책의 강화, 뉴딜 정책기의 자유주의로의 전환, 그리고 1970년대 이후 수정주의적 자유주의로의 변화 과정을 구체적으로 살펴본다. 그리고 현재 미국의 무역정책을 상계 관세와 반덤핑 관세를 중심으로 알아보고, 최근 쟁점과 사안에 대해 고찰해 본다.

제 4장~제 6장은 본 연구의 핵심적인 부분인 데이터 분석에 관한 개요이다. 데이터 분석의 목적과 활용된 데이터에 대해 설명을 하고, 제 5장과 제 6장에서 각각 군집분석과 의사결정나무 분석, 그리고 순환형 신경망 방법론에

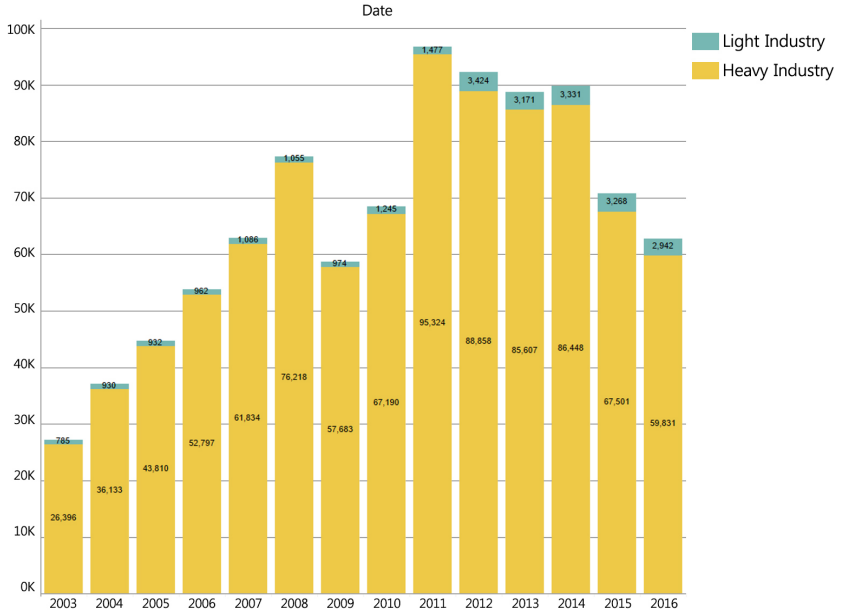
대해 알아본 후 이들 방법론을 활용해 분석을 수행한 결과를 살펴본다. 군집분석은 분류되어 있지 않은 데이터를 상이한 특성을 갖는 '군집'으로 분류하고 비교하는 데 효과적인 비지도 학습(unsupervised learning) 방식으로, 각 연도를 상계 관세와 반덤핑 관세 부과 추세를 기준으로 군집화하고 보호주의 무역 기조의 존재 여부에 따라 비교하기는 데에 좋은 도구가 될 수 있다. 한편 순환형 신경망은 순차형 데이터(sequential data)에 대한 높은 예측력을 보이는 기계학습 모델로, 과거의 추세가 현재와 미래에 영향을 미치는 무역 및 기업 경영 데이터 분석에 적합하다고 할 수 있다. 다만 그러한 예측하는 과정을 '설명'하고 어떠한 변수가 주요한 영향력을 갖는지 밝혀내기에는 적합하지 않다는 점을 감안하여 본 연구에서는 결과해석과 '설명'이 쉬운 의사결정나무 분석 방법론으로 한계점을 보완하고자 한다. 마지막으로, 제 7장은 본 연구의 결론으로 전반적인 내용 요약과 함께 연구의 시사점 및 한계점을 제시한다.

2. 현황 및 추이

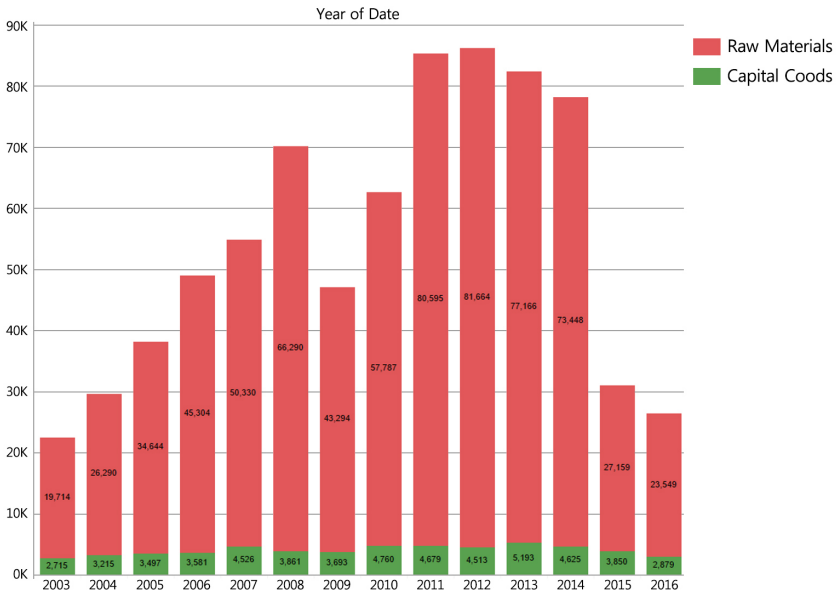
2.1 전반적인 수출 추이

울산지역의 주력산업은 자동차, 조선, 석유화학, 정유 등 중화학공업 중심으로, 울산지역 수출의 대부분은 중화학 제품이며(<Figure 1> 참조), 수입의 대부분은 공정의 원료가 되는 원자재이다(<Figure 2> 참조).

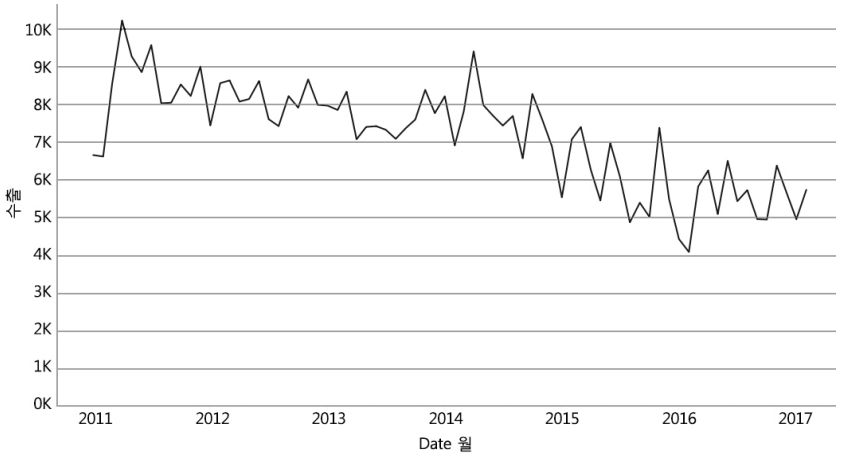
우리나라 제조업의 전반적인 발전과 함께 울산지역의 중화학 제품 수출액도 2000년 이후 꾸준히 성장세를 이어왔다. 하지만 2011년($95,324,000,000)을 정점으로 조선과 정유 등



〈Figure 1〉 Export Statistics of Ulsan City[6]



〈Figure 2〉 Import Statistics of Ulsan City[6]



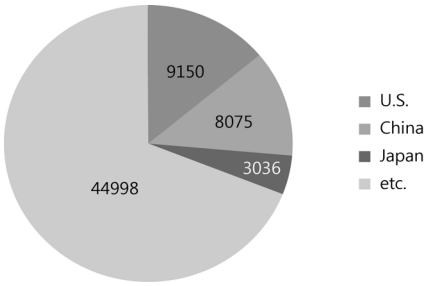
〈Figure 3〉 Monthly Exports of Ulsan City[6]

이 동반 침체를 겪으면서 2017년 현재까지 지속적으로 수출액이 감소하고 있다(<Figure 1> 참조). 이와 더불어, 원자재 수입액도 지속적인 감소세를 보이고 있으며(<Figure 2> 참조), 월별 수출액도 증감을 보이고 있으나 전반적으로 상당한 감소 추세를 나타내고 있다(<Figure 3> 참조).

2017년 초 수출액이 전년도 대비 40% 가량 증가하며 회복세를 보였으나[55], 국제유가 상승 등으로 인한 일시적인 현상으로 장기적으로 오름세를 이어가기는 힘들 것이라는 분석도 제시되고 있는 상황이다[53].

2.2 국가별 수출 추이

울산지역 수출의 가장 큰 비중을 차지하는 국가는 미국, 중국, 그리고 일본이다. 2016년 울산지역에서 3개 국가에 대한 수출액은 전체 수출액의 20,261백만 달러로 전체 수출액의 1/3정도를 차지하고 있다(<Figure 4> 참조). 특히 미국은 가장 큰 비중을 차지하고 있는데, 중국과 일본이 우리나라에서 중간 단계의 제품을 구입해 완제품을 미국에 수출하는 '간접 수출' 효과도 무시할 수 없으므로 미국과의 관계는 울산지역의 수출과 경제에 막대한 영향을 미친다고 할 수 있다[6].



〈Figure 4〉 Export Scales of Ulsan City〔6〕

그러나 미국 등 주요 무역국을 대상으로 한 울산지역 산업의 수출 전망은 밝지만은 않다. 2016년 영국이 유럽연합(EU)에서 탈퇴하는 '브렉시트'[40] 이후 자국의 이익을 보호하려고 하는 보호주의(Protectionism) 기조가 전 세계적으로 강화되면서 울산지역 주력산업 수출 증대의 제약요인으로 작용할 전망이다[44].

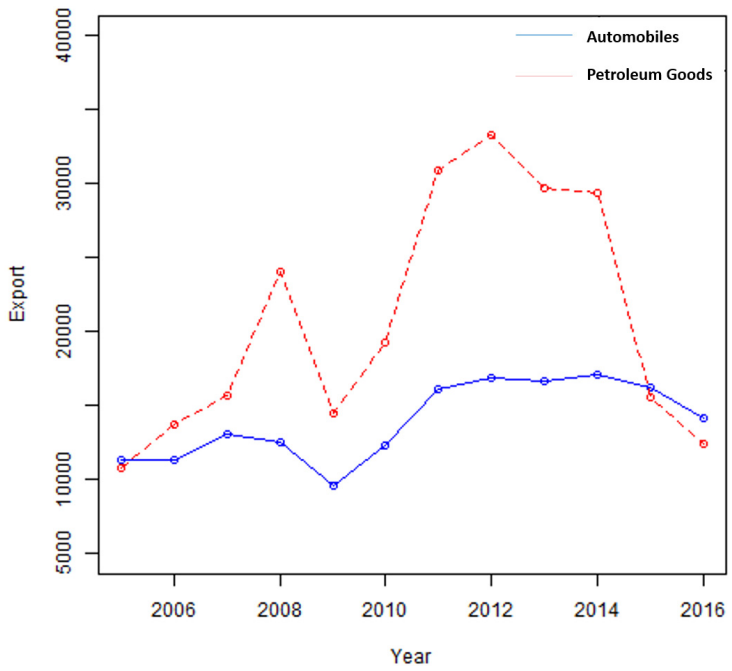
특히 2017년 초 도널드 트럼프(Donald Trump) 미국 대통령의 취임과 한반도 사드(THAAD, 고고도미사일방어체계) 배치에 따른 중국의 보복무역으로 인해 울산지역의 수출 기업들이 예전보다 심각한 난조를 겪을 것으로 예상되고 있다[43, 51, 52].

2.3 품목별 수출 추이

앞서 언급하였듯이 울산의 주력산업은 중화학공업 중심으로 석유 제품, 석유화학, 자동차, 자동차부품, 그리고 조선이 울산지역의 5대 수출품목이다. 울산지역의 5대 수출품목의 구체적인 연도별 수출액은 <Table 1>과 같다.

〈Table 1〉 Export Scales of Ulsan City〔5〕

Year	Petroleum Products	Petroleum Chemicals	Shipbuilding	Automobiles	Automobile Components	Total
2005	10,735	7,118	6,404	11,368	2,037	37,662
2006	13,668	8,371	7,833	11,373	2,662	43,907
2007	15,715	10,247	9,592	13,112	2,828	51,494
2008	24,019	11,355	13,109	12,558	3,153	64,194
2009	14,460	9,054	11,949	9,580	1,760	46,803
2010	19,260	10,899	9,998	12,332	2,491	54,980
2011	30,918	14,803	13,640	16,084	3,262	78,707
2012	33,346	13,685	7,649	16,855	3,671	75,206
2013	29,698	13,801	8,219	16,675	3,860	72,253
2014	39,396	12,754	11,707	17,090	3,027	73,974
2015	15,577	7,362	13,902	16,219	2,348	55,408
2016	12,424	7,441	10,259	14,190	2,706	47,020
Sum	**249,216**	**126,890**	**124,261**	**167,436**	**33,805**	**701,608**



〈Figure 5〉 Export Scales of Automobile and Petroleum Industries[5]

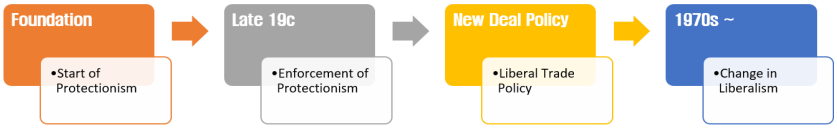
5대 수출 품목의 수출액 추이는 울산지역의 전체 수출액(<Figure 1> 참조)과 유사한 증감 추세를 보이고 있다. 울산지역의 5대 수출 품목은 2011년에 정점을 찍은 후($78,707,000,000) 꾸준히 하락세를 그려 2016년에는 2009년도($46,803,000,000)와 비슷한 수준($47,020,000,000)으로 회귀하였다.

2005년 이후 지난해까지 누적 수출액을 보면 석유 제품($249,216,000,000)이 가장 큰 비중을 차지하고 자동차 부품($33,805,000,000)이 가장 적은 비중을 나타내고 있다. 한편, 석유 제품은 연도별 변동성도 가장 크며 자동차가 변동성이 가장 적은 제품임을 알 수 있다(<Figure 5> 참조). 이는 석유 제품이 미국 무역정책의 변화와 같은 외부 효과에 영향을 받을 가능성이 높다는 것을 시사한다고 할 수 있다.

3. 문헌 연구

3.1 미국 무역정책의 변동

미국 무역정책의 발전 과정은 크게 네 단계로 나누어 볼 수 있다(<Figure 6> 참조). 건국 초기 재정 수입을 증대시키고 세계적인 중상주의 추세에 발맞추기 위해 미국 연방 정부는 관세율을 높게 설정하였다. 이와 함께, 자국 내 제조업 진흥(promotion)과 기반시설 확충(internal improvement) 등을 통한 신생국(New Nation)으로서의 경제적 자립을 달성하기 위한 정치적 목적으로 보호무역주의 기조는 더욱 강화되었다. ≪가증스러운 관세(tariff of abomination)≫라고 불리던 1828년의 관세법이 이 시기의 보호주의 기조를 상징하는 대표적인 사례라고 할 수 있다[3, 19].



〈Figure 6〉 Historical Changes in U.S. Trade Policy

1800년대 후반에 기존의 보호주의 정책이 도전 받으면서 일시적으로 자유무역주의 정책이 급부상하게 된다. 대표적으로, 우드로 윌슨(Woodrow Wilson) 대통령은 무역 자유화 정책과 함께 1913년 언더우드 관세법(Underwood Tariff Act)을 제정해 기존의 관세를 다소 완화하였다. 그러나 1918년 공화당이 의회에서 주도권을 잡으면서 다시 보호주의 기조가 강화되었다. 이후 포드니-매컴버 관세법(Fordney-McCumber Tariff Act)과 스무트-홀리 관세법(Smoot-Hawly Tariff Act) 등이 제정되면서 미국의 보호주의 체제는 날이 갈수록 공고해졌다[3, 29, 41].

대공황에 대응하여 실시한 프랭클린 루즈벨트(Frankling Roosevelt) 대통령의 초기 뉴딜(New Deal) 정책은 공화당 연합이 형성한 보호주의 기조를 유지하는 것이었다. 하지만 내부적인 모순과 갈등을 극복하고자 한 ≪제2단계≫ 뉴딜정책 연합은 자유무역에 대한 요구를 수용하였다. 따라서, 1934년 호혜통상법(Reciprocal Trade Agreements Act; RTAA)이 제정되며 미국 무역정책이 자유주의로 방향을 선회하게 된다. 나아가, 제 2차 세계 대전이 종료되면서 냉전 시대를 맞아 미국의 자유주의 정책이 전성기를 맞게 된다[3, 15, 17].

이어 1950년대와 1960년대에 이르러 세계 경제에 미국의 패권주의가 강화되면서 자유무역주의 기조가 공고해졌다. 하지만 1970~1980년대에 들어 미국의 대외 무역 적자가 기하급수적으로 늘어나면서 보호주의로의 회귀 주장이 제기되었다. 그렇지만 한편으로는 미국경제의 수출 의존도도 지속적으로 증대되어 완전한 보호주의로의 복귀는 현실적으로 불가능한 상황이었다. 특히 반도체, 항공기 등의 고부가가치 첨단산업의 무역의존도가 높아지면서 보호주의로의 회귀는 미국경제 전반에 치명적인 영향을 미칠 수도 있는 상황에 직면하였다(<Table 2> 참조). 이에 수정주의적 자유주의(revisionist liberalism)로 분류되는 새로운 요구들이 등장하였다. 미국의 첨단 산업의 약화 요인이 외국 정부들의 자국 산업을 보호하기 위한 보호주의 무역 기조에 있으며 공정무역이나 전략적 무역정책 등을 통해 이러한 문제를 시정해야 한다는 것이 이러한 주장들의 핵심내용이었다[3, 12, 33].

〈Table 2〉 Level of Dependence on Exports of United States Economy in 1981[3]

Product	Proportion	Product	Proportion
Aircraft Components	45%	Wheat	77%
Semiconductor	42%	Tractor Components	63%
Science Equipments	40%	Raw Cotton	52%
Construction Equipments	38%	Soybean	50%
Automobile Components	30%	Agriculture Equipments	42%
Electricity Measures	28%	Corn	39%

3.2 현재의 미국 무역정책

2000년대에 들어 보호주의적 무역정책은 자국의 이익을 다른 나라의 불공정행위로부터 보호하고자 하는 상호주의에 입각한 ≪공정 무역(fair trade)≫ 정책으로 지속되어 왔다. 특히 2017년 초 도널드 트럼프 제 45대 대통령이 취임하면서 미 행정부는 교역상대국의 ≪불공정 무역관행(unfair trade practices)≫에 엄중하게 대처할 것을 발표했다. '2017년 대통령 통상정책의제'에서 ≪자유무역≫ 또는 ≪자유시장≫이라는 단어가 8번 등장하는데 불과한 반면, ≪공정≫ 혹은 ≪불공정≫이라는 단어는 28번이나 언급되었다[7, 10].

1950년대 이후 자유무역 기조가 상대적으로 강해졌음에도 불구하고 아직 미국의 무역정책은 수많은 규제로 보호무역의 색깔을 강하게 띠고 있다. 국제무역 연구기관인 Global Trade Alert에 따르면, 미국은 수출 보조(Export subsidy), 수입 금지(Import ban), 구제 금융(Bail-out) 등 보호주의적 성격의 규제 장치를 450여 개나 만들어 자국 산업을 보호하기 위해 노력하고 있다. 이에 반해 브라질이나 중국, 유럽, 일본 등은 비슷한 성격의 규제 장치의 개수가 150개에 미치지 못한다. 동일 기관에 따르면, 2017년 2월 보호무역 관련 누적 집계 사건 수는 미국이 993건, 아세안 국가가 316건, 중국이 154건에 이르고 있다[8].

대표적인 보호주의적 공정 무역정책의 규제가 바로 관세 제도(Tariff System)이다. 미 무역위원회(United States International Trade Commission)는 불공정 무역으로 인해 무역 질서를 해치는 경우 상계 관세(Countervailing Duties)와 반덤핑 관세(Antidumping Duties)를 징수하고 있다. 불공정한 수입 제품으로 인해 미국 기업이 부당한 피해를 보았다는 진정서(Petition)를 제출하면, 무역위원회가 이를 조사하고 불공정하다고 판단되는 경우 상계 관세나 반덤핑 관세를 부과한다[36, 56].

상계 관세는 수출국에서 기업에 지급하는 보조금으로 인해 미국에 수출하는 가격이 내려가는 경우 이러한 가격차를 상쇄하기 위해 부과하는 특별 관세이다. 반덤핑 관세는 수출기업이 정상 가격보다 심각하게 낮은 가격으로 미국에 수출하여 미국 기업이 받는 피해를 줄이기 위해 부과되는 특별 관세이다[38]. 미 무역위원회에 따르면, 1980년부터 2008년 까지 접수된 상계 관세 및 반덤핑 관세 관련 사건은 1,632건이며 반덤핑 관세는 연평균 40건, 상계 관세는 연평균 16건으로 반덤핑 관세 사례가 상계 관세 사례에 비해 두 배 이상 많았다[56].

미 무역위원회 홈페이지(https://www.usitc.gov)에서는 반덤핑 및 상계 관세 관련하여 현재 진행 중인(active) 사건과 완료된(final) 사건 목록과 상세 설명을 찾아볼 수 있다(<Figure 7> 참조). <Table 3>는 현재 진행 중인, 혹은 완료된 상계 관세 및 반덤핑 관세 사건의 예시이다.

실제로 트럼프 행정부가 들어선 이후 우리나라 기업들이 미국에서 강한 규제의대상이 되고 있다. 특히, 무역특혜연장법 504조에 의해 많은 기업들이 ≪징벌적 규제≫의 대상이 되고 있다[45]. 대표적인 사례로, 미국 정부는 지난 3월 현대중공업이 생산한 대형 변압기(transformer) 수출 제품에 예비판정(3.09%)의 약 20배에 달하는 61%의 높은 반덤핑 관세를 확정하였다[22, 37]. 이에 현대중공업은 생산 시설을 미국 공장으로 이전해 ≪관세 폭탄≫을 피하고자 할 정도로 심각성을 느끼고 있다[2].

〈Figure 7〉 Antidumping and Countervailing Related Cases[1)]

〈Table 3〉 Instances of Antidumping and Countervailing Related Cases[2)]

State	Title	Code	Nation	Phase
Active	Dioctyl Terephthalate(DOHP) from Korea	731-TA-1330	Korea	Final Phase
Active	Carbon and Alloy Seamless Standard, Line, and Pressure Pipe from Japan and Romania	No. 731-TA-847 and 849	Japan, Romania	Third Review
Completed	Tetrafluoroethane(R-134a) from China	No. 731-TA-1313	China	Preliminary Phase

3.3 미국 무역정책의 쟁점과 사안

올해 초 현대중공업의 사례만 보아도 미국 무역정책의 변화가 우리나라 기업의 경영성과에 막대한 영향을 미치는 것을 알 수 있다. 나아가, **2. 울산지역 주력산업의 해외 수출현황과 추이**에서 살펴보았듯이, 미국 수출에 높은 의존도를 보이고 있는 중화학 공업 위주의 울산지역 주력산업은 미국의 통상 정책의 변동에 큰 영향을 받을 수 밖에 없다.

이러한 점을 감안하여 미국 무역위원회에서 엄중하게 다루고 있는 반덤핑 관세 및 상계 관세를 중심으로 미국 무역정책 상의 변화가 수출 기업의 경영성과에 미치는 영향에 대해 살펴보려고 한다. 나아가, 미국 무역정책의 변동에 따른 울산지역의 주력산업 경영성과의 변동 방향을 예측할 수 있는 방안을 찾아보고자 한다.

1) https://www.usitc.gov/trade_remedy/731_ad_701_cvd/investigations.htm.
2) https://www.usitc.gov/trade_remedy/731_ad_701_cvd/investigations.htm.

4. 분석 개요

4.1 분석 목적

본 연구의 분석 목적은 크게 두 가지이다. 먼저 반덤핑 관세와 상계 관세를 중심으로, 미국 무역정책의 변화가 울산지역 주력산업의 경영성과에 미치는 영향 여부를 검증한다. 또한 미국의 무역정책 상의 변화가 울산지역 주력산업의 경영성과에 미치는 영향을 예측하고, 나아가 그러한 영향의 정도를 수치화 한다.

위의 목적에 부합하는 객관적이고 현실적인 분석을 수행하기 위하여 우리는 데이터 기반(Data-driven) 검증 및 예측 기법을 채택하였다. 활용 가능한 데이터의 양과 종류가 많아지고 데이터를 분석할 수 있는 기술이 발달하면서 데이터에 기반한 의사결정의 효과성과 중요성이 갈수록 증대되고 있다[26, 31, 41].

본 연구의 접근법이 기존의 계량경제학 및 기초 통계학에 기반한 경제학적인 분석 방식과 갖는 가장 큰 차별점은 보다 다차원적이고 변수 간에 내부적으로 상호작용하는 모델을 통해 선형적으로 설명되지 않는 복잡다단한 현실 세계의 현상을 설명 및 예측할 수 있다는 것이다. 회귀분석에 기반한 경제학적 모델들은 수학적으로 정교하고 제한된 변수를 통해 현상에 대한 설명을 이해하기 쉽다는 장점을 가지고 있지만, 복잡한 현실을 반영하지 못하는 지나치게 단순한 모델 구조로 인해 현상에 대한 설명력이 현저하게 떨어진다는 한계점이 존재한다. 뿐만 아니라, 현상에 대한 '설명(description)'에 지나치게 초점을 맞추다 보니 예측력(prediction power)이 현실의 문제 해결에 적용 불가능할 정도로 떨어진다는 약점이 있다[49]. 예를 들어, 지난 2007년에 많은 경제학자들이 확신을 가지고 주택 가격 거품(housing bubble)이 없다고 주장하였지만, 현실에서는 미국 주택 시장이 붕괴함에 따라 세계적인 규모의 경제 위기가 찾아왔다[48].

나아가, 본 연구에서 부분적으로 적용하는 딥 러닝(deep learning) 기반 학습 모델의 경우 기존의 기계학습 모델에서 연구자가 수행하였던 특징 추출(feature extraction)까지 기계가 대신 수행해줄 만큼 강력하다. 데이터가 많아지고 순환형 신경망(recurrent neural networks), 합성곱 신경망(convolutional neural networks), 비선형 활성 함수(nonlinearities), 배치 정규화(batch normalization) 등 관련 모델과 테크닉들이 발달하면서 접목 분야가 광범위해지고 있다. 자율주행 자동차(autonomous vehicle), 이미지 처리(image processing), 비디오 자막 생성(caption generation) 등 딥 러닝 기술의 적용 분야는 갈수록 다양해지고 깊어지고 있다.

4.2 데이터

미국의 무역정책과 관련한 자료로는 미 무역위원회에서 제공하는 'Import Injury Investigations Case Statistics'[56]를 참고하였다. 이는 1980년부터 2008년까지 29년 동안 발생한 수입상해사건(Import injury cases)을 연도별로 정리해 놓은 데이터이다. 그 중에서도 우리는 구체적으로 미국 연방 조세법(the Tariff Act of 1930)의 타이틀 Ⅶ(Title Ⅶ)에 저촉되는 반덤핑 관세와 상계 관세 부과 사례와 관련된 통계 자료를 활용하였다(<Table 4> 참조).

〈Table 4〉 Statistics Regarding Antidumping and Countervailing Duties by USITC (Unites States International Trade Commission 2010)

Fiscal Year	Number of Cases	Result			Import Value
		Affirmative	Negative	Terminated	
2008	26	23	1	2	2,222,355
2007	37	25	11	1	1,986,708
2006	8	2	3	3	1,912,478
2005	12	8	4	0	1,688,847
2004	39	21	9	9	1,469,305
2003	41	16	14	11	1,288,523
2002	39	13	22	4	1,176,948
2001	116	55	48	13	1,233,117
2000	36	19	15	2	1,225,314
1999	67	26	29	12	1,028,588
1998	44	29	12	3	944,508
1997	29	15	11	3	891,227
1996	14	9	2	3	814,621
1995	20	11	6	3	775,640
1994	70	31	32	7	676,563
1993	37	12	9	16	601,983
1992	133	56	72	5	550,897
1991	74	20	43	11	523,659
1990	25	16	6	3	518,210
1989	32	19	9	4	497,086
1988	48	24	17	7	463,463
1987	44	21	16	7	419,218
1986	89	47	21	21	393,745
1985	120	34	32	54	362,073
1984	80	24	21	35	340,567
1983	42	15	16	11	265,733
1982	178	30	78	70	269,226
1981	32	5	5	22	407,377
1980	100	11	70	19	256,047
Sum	**1,632**	**637**	**634**	**361**	**25,204,027**

사례 수는 해당 회계연도 중에 제기된 모든 반덤핑 관세 및 상계 관세 사건의 수이고, 모든 사례의 판결 결과는 승인(Affirmative), 반려(Negative), 혹은 종결(Terminated) 세 가지 중 하나이다. 사례가 승인될 경우 반덤핑 및 상계 관세가 특정 품목에 대해 부과되고, 반려될 경우에는 부과되지 않는다. 한편, 사건 신청인이 사건을 철회하거나 당국이 이를 유예할 경우 사건이 종결된다. 마지막으로, 수입 금액은 사례와 관련된 수입 항목의 달러 가치를 나타낸다[56].

한편, 울산지역 주력산업의 경영성과와 관련된 데이터는 한국은행의 기업경영분석 통계를 활용하였다. 매년 발행되는 한국은행 기업경영분석 보고서를 통해 석유정제, 조선, 자동차, 그리고 자동차 부품 제조 산업의 경영성과 데이터를 얻을 수 있었다. 구체적으로, 각 산업의 연도별 매출총이익(Gross profit on sales)과 영업이익(Business profits) 데이터를 수집하였다. 기업의 손익계산서 상의 매출총이익과 영업이익은 경상이익이나 당기순이익에 비해 기업의 실질적인 성과와 향후 성장성을 측정할 수 있는 적절한 지표로 알려져 있다[9].

5. 방법론

5.1 군집분석(Clustering Analysis)

우선 첫 번째 분석 목적을 달성하기 위하여, 미국 무역정책 데이터에 대해 연도별로 군집분석(Cluster Anaylsis)을 수행하였다. 군집분석은 기계학습(Machine learning) 방법론 중 비지도학습(Unsupervised learning)의 일종으로, 전체 관측된 데이터를 내부적으로 유사한 군집(clusters)들로 분할(partition)하는 분석 방법이다.

군집분석의 대표적인 적용 사례로 효과적인 마케팅을 위해 유사한 고객끼리 묶어 분류하는 시장 세분화(Market segmentation)를 들 수 있다. 군집분석의 방법에는 크게 위계적 군집화(Hierarchical clustering)와 k-평균 군집화(k-means clustering)의 두 가지 방식이 있는데, 본 연구에서는 미리 군집의 개수를 설정해주고 중심점을 바꿔가며 적절한 군집을 찾아가는 k-평균 군집화 방식을 사용하였다[16, 45].

본 연구에서 군집분석을 활용한 이유는 보호주의 무역 기조가 강했던 연도와 그렇지 않았던 연도를 과거 데이터를 가지고 객관적으로 분류하기 위해서이다. 각 연도별 추세를 대변하는 대표값이 존재하지 않으므로 지도학습 기법 중 분류(classification) 기법을 활용할 수 없고 다차원의 데이터이므로 사람이 각 차원에 알맞은 기준을 일일이 손으로 찾아낼 수 없으므로 비지도학습 기법인 군집분석을 활용하였다.

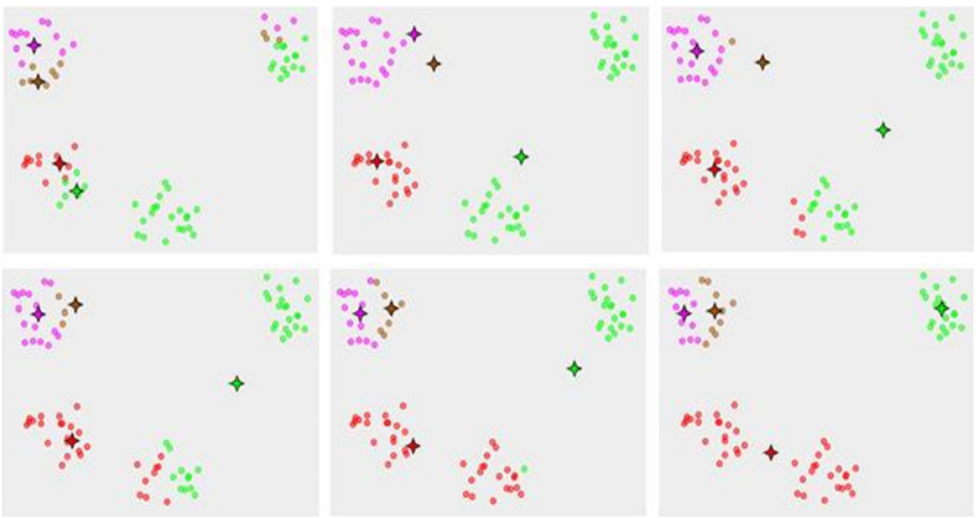
n개의 관측치 $\{e_1, e_2, \cdots, e_n\}$이 있을 때, k-평균 군집화의 최종 목적은 d차원의 벡터 공간상에 존재하는 각 관측치 e_1를 k개의 집합 $P=\{P_1, P_2, \cdots, P_k\}$로 분류하되 군집 내부의 분산(variance)이 최소화 되도록 하는 것이다. 일반적으로 군집 내부의 분산은 군집 내의 원소들 간의 거리(Euclidean distance)로 계산된다. 즉, 군집분석의 목적은 아래 군집내부제곱합(WCSS, Within-cluster sum of squares)을 최소화하는 것이다[1, 45].

$$WCSS = \sum_{P_i \in P} \sum_{j=1,\cdots,d} \sum_{a,b \in P_i} (a_{ij} - b_{ij})^2$$

$$= \sum_{P_i \in P} \sum_{j=1,\cdots,d} 2|P_i| \sum_{a,b \in P_i} (a_{ij} - \mu_{ij})^2$$

where μ_{ij} is the mean coordinate of cluster i in dimension j (i.e., centroid)

즉, 군집내부제곱합은 각 집합의 중심점(centroid)으로부터 그 집합에 속한 관측치까지의 거리를 제곱해 합한 값과 같으며, k-평균 군집분석 알고리즘은 처음에 초기 중심점(μ_i)



〈Figure 8〉 How the *k*-Means Algorithm Finds Optimal Clusters〔36〕

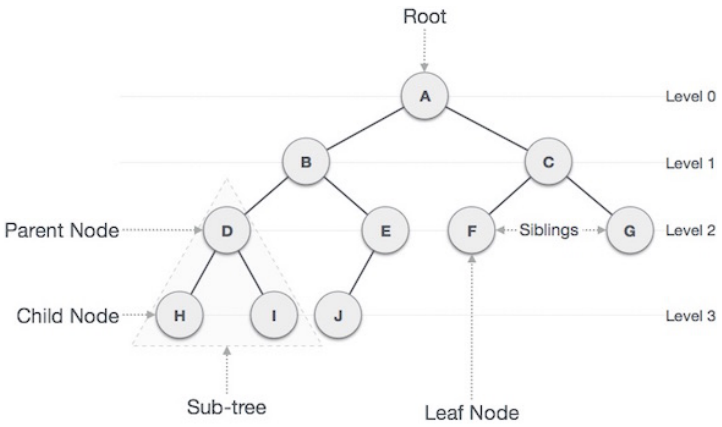
을 설정하고 각 관측치 가장 가까운 군집을 배당하고 중심점을 각 군집의 관측치들의 무게중심으로 재조정하는 과정을 유한 번 반복해 최적의 군집을 찾아낸다[21, 23, 45]. <Figure 8>은 *k*-평균 알고리즘이 최적 군집을 찾아가는 과정을 시각화한 것이다[36].

본 연구에서는 *k*-평균 군집분석 알고리즘을 활용하여, 1980년부터 2008년까지 각 연도를 수입 상해 사건 관련 통계(<Table 4> 참조)를 기반으로 두 개의 군집으로 분류하였다. 연도별 군집화의 주목적은 수입 상해 사건 발생 경향이 비슷한 두 군집으로 나누어, 두 군집 기간 동안에 울산지역 주력산업의 경영성과에 유의미한 차이가 있는지를 알아보기 위함이다. 즉, 두 군집 사이에 울산지역의 경영성과(매출총이익 및 영업이익 증가율)를 비교하여 미국 무역정책의 보호주의적인 기조가 울산지역 주력산업(자동차/자동차 부품/석유정제/조선)의 경영성과에 어느 정도의 영향을 미치는지를 살펴보도록 하였다.

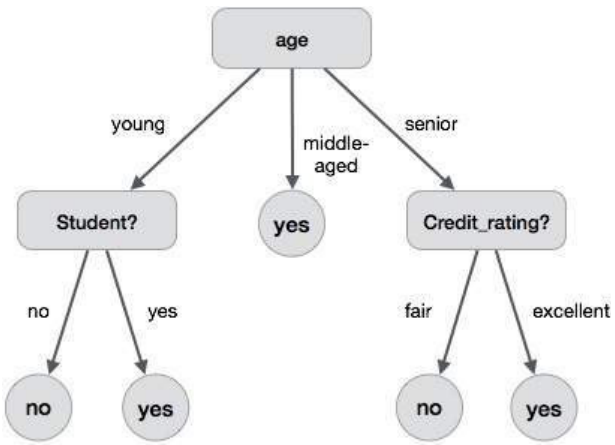
5.2 의사결정나무 분석(Decision Tree Analysis)

군집분석에 이어, 미국 무역정책의 변화가 울산지역 주력산업에 미치는 영향을 예측하기 위하여 의사결정나무 분석(Decision Tree Analysis)을 수행하였다. 의사결정나무는 분류(classification)와 예측(prediction)을 위해 대중적으로 널리 쓰이는 기계학습 알고리즘 중 하나로 직관적이면서도 결과를 해석하기 쉽다는 장점이 있다. 의사결정나무 분석 결과는 여러 개의 단순한 분기 조건인 *if-then* 조건문(*If X, then Y*)들로 이루어져 있으며, 시각화(visualize)하고 설명하기 쉽다[31, 45].

본 연구에서는 순차형 데이터를 다루는 데 용이하고 상당히 높은 예측력을 보이고 있는 순환형 신경망 모델과 함께 시각화와 결과를 설명하고 해석하기 쉬운 의사결정나무 분석을 동시에 활용해 각각의 장점을 극대화하고 단점을 보완하고자 하였다. 나아가, 의사결정나무



〈Figure 9〉 Structure of a General Binary Tree[53]



〈Figure 10〉 Decision Tree to Predict a Customer's Purchase[52]

는 각 노드의 내용을 통해 중요한 변수(feature)를 밝혀낼 수 있다는 장점도 있으므로 보다 폭넓은 결과 해석을 위해 도입하였다.

의사결정나무는 트리 자료구조(Tree data structure)를 기반으로 하며, 특히 이진 트리(Binary tree)와 유사한 형태를 보인다. 이진 트리는 노드(node)와 엣지(edge)로 이루어져 있는 자료 구조이며, 가장 위에 있는 노드를 뿌리(root)라고 지칭한다. 모든 이진 트리에는 하나의 뿌리 노드만 존재한다. 특정 노드에 연결된 노드 중 위쪽에 위치한 노드를 부모 노드(parent node)라고 하며, 아래쪽에 위치한 노드를 자식 노드(child node)라고 한다. 자식 노드를 하나라도 가지고 있는 노드를 내부 노드(internal node)라고 하며 자식 노드를 가지고 있는 않는 노드를 단말 노드(leaf node)라고 한다. <Figure 9>는 일반적인 이진 트리의 구조이다[44, 49].

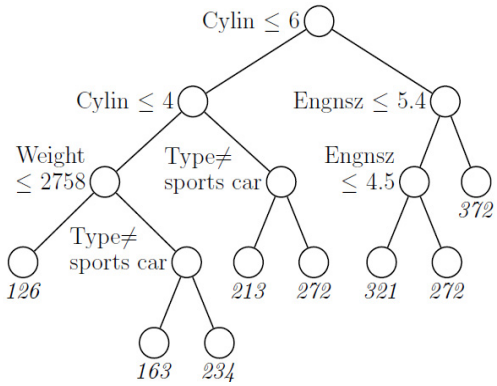
<Figure 10>은 고객의 컴퓨터 구입 여부를 예측하기 위한 간단한 형태의 의사결정나무이다[51]. 전체적인 형태는 앞서 설명한 이진 트

〈Table 5〉 All Cases of the Decision Tree in 〈Figure 10〉

	Age	Student	Credit_rating	Classification
1	young	O	N/A	YES
2		X		NO
3	middle-aged	N/A	N/A	YES
4	senior	N/A	excellent	YES
5			fair	NO

리와 거의 유사한, 나무를 거꾸로 뒤집어 놓은 형상이라고 생각하면 이해가 쉬우며, 가장 위의 뿌리 노드로부터 시작해 각각의 내부 노드마다 *if- then* 논리를 적용해 가장 아래의 단말 노드까지 내려오면 예측 혹은 분류가 완료된다. 예를 들어, <Figure 10>의 의사결정나무에서 뿌리 노드에서는 나이(age)를 가지고 조건문을 형성한다. 만약 나이가 어릴(young) 경우, 가장 왼쪽의 노드(Student?)로 내려가고, 중년(middle-aged)일 경우 바로 중간의 단말 노드로 내려가 컴퓨터를 구입할 것이라고 예측하며, 노년(senior)일 경우 가장 오른쪽의 노드(Credit_rating?)로 내려가 새로운 *if-then* 조건문을 시작한다. <Table 5>는 <Figure 10>의 의사결정나무의 모든 경우의 수를 고려해 최종 분류 결과를 나타낸 표이다.

본 연구에서는, 예측을 위해 회귀나무(Regression tree)를 활용하였다. 회귀나무는 <Figure 10>과 같은 분류나무(Classification tree)와 거의 유사하지만, 단말 노드의 값이 이산형(discrete)이 아닌 연속형 변수(continuous variable)에 속한다는 차이점이 있다[33]. <Figure 11>은 회귀나무의 예시이다. <Figure 11>의 회귀나무의 경우, 단말 노드의 값이 126, 163, 213 등과 같이 정수형(Integer)에 속하는 경우이다[33].



〈Figure 11〉 Example of a Regression Tree[33]

회귀나무의 경우 오차를 일반적으로 평균제곱합(MSE; Mean squared error)으로 측정한다. 평균제곱합은 각각의 예측치($\hat{y}_i$)에서 실제 관측치(y_i)를 뺀 값을 제곱하여 평균한 값이다. 총 n개의 예측을 시도한 경우 평균제곱합은 아래와 같다[47].

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

본 연구에서는 반덤핑 및 상계 관세 사례 데이터와 울산지역 주력산업 경영성과 데이터를 기반으로 회귀나무를 생성하여, 미국의 보호주의적 무역 행태에 따라 울산지역의 주력산업의 경영성과 예측을 시도하였다.

5.3 순환형 신경망(RNN; Recurrent Neural Networks)

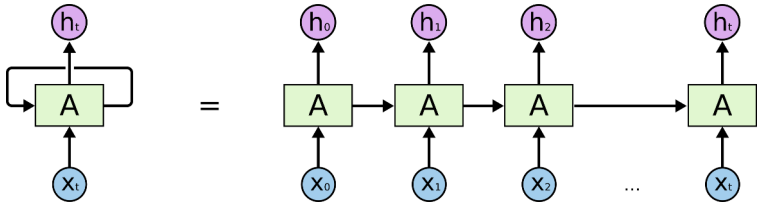
기업 경영성과와 업종별 수출액 등은 주식과 같이 과거의 상태가 현재의 상태에 지속적인 영향을 미치는 순차적인 형태의 데이터(sequential data)라고 할 수 있다. 하지만 선형회귀분석(linear regression), 의사결정나무, 지지벡터머신(support vector machines) 등의 일반적인 기계학습 알고리즘은 이러한 형태의 데이터를 처리하는 데 공히 어느 정도 한계점을 가진다고 할 수 있다.

순환형 신경망은 내부에 순환하는 구조를 내재하고 있는 네트워크로, 그러한 구조는 ≪정보가 지속되는 것을 돕는다≫. <Figure 12>는 일반적인 순환형 신경망 구조를 시각화한 것으로, 각 네트워크는 다음 네트워크에게 정보를 전달하며 사슬과 같은 구조를 형성한다[11, 20].

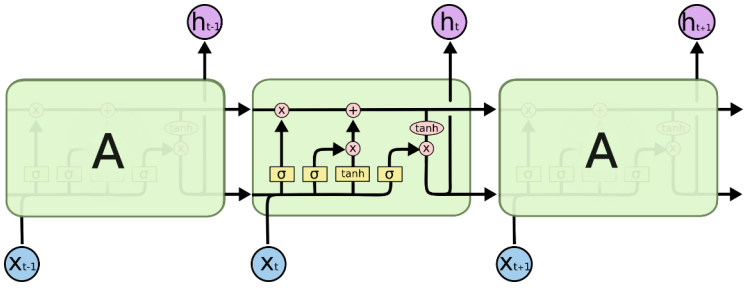
예를 들어, ≪the clouds are in the sky≫라는 영어 문장을 학습시킨 후 검증을 위해 ≪the clouds are in the≫까지만 주더라도 순환형 신경망은 다음으로 올 단어가 ≪sky≫임을 쉽게 예측할 수 있다[11].

그렇지만 보다 문맥이 확장되어야 하는 경우에는 일반적은 순환형 신경망은 적절하지 않다. 예를 들어 ≪나는 프랑스에서 자랐습니다… 나는 유창한 프랑스어를 할 수 있습니다.≫라는 문장에서 ≪프랑스어≫라는 단어를 예측하려 한다면 문장의 처음에 있는 프랑스라는 문맥이 필요하다[11].

이러한 기억의 ≪장기 의존성≫을 가능케 한 신경망 모델이 장단기 기억 네트워크(LSTM; Long short term memory networks)이다. 장단기 기억 네트워크는 일반적인 순환형 신경망보다 복잡한 내부 구조를 형성함으로써 오랫동안 정보를 기억하는 것을 가능케 한다. <Figure 13>은 일반적인 장단기 기억 네트워크의 구조이다[11, 24].



〈Figure 12〉 Structure of a Vanilla Recurrent Neural Nets[20]



〈Figure 13〉 Structure of a Long Short-Term Memory Cell[20]

이에 우리는 장단기 기억 네트워크 모델을 활용하여 최근 5년 간의 기업 경영 및 수출 관련 데이터를 통해 울산지역의 기업 경영성과, 구체적으로는 매출총이익을 예측해 보고자 하였다. 장단기 기억 네트워크 모델은 순차형 데이터를 학습해 예측하는 데 놀라울 정도로 높은 정확도를 보이는 모델로 향후 울산지역의 기업 경영성과를 예측하는 데에도 효율적인 모델이 될 것이라 생각해 도입하였다.

장단기 기억 네트워크를 포함한 대부분의 신경망 모델에서 중요한 것은 모델의 크기를 설정하는 셀 크기(cell size), 데이터 학습 횟수인 에포크(epochs), 모델의 과적합을 막기 위한 드랍아웃 비율(dropout rate), 한 번에 몇 개의 데이터 인스턴스를 처리하는 지를 나타내는 배치의 크기(batch size) 등의 하이퍼 파라미터(hyper-parameter)를 적절히 설정하는 것이다. 이러한 하이퍼 파라미터는 연구자가 임의로 설정할 수 있으나, 세부 설정에 따라 모델의 효능이 크게 달라질 수 있으므로 주의를 기울여야 한다. 우리는 다양한 경우의 수를 테스트해보기 위해 연구자가 설정한 모든 경우의 수를 각각의 모델을 세워 교차검증(cross validation)을 통해 앞에서 설명한 평균제곱오차(mean squared error)를 계산해 보았다.

6. 결 과

6.1 분석 결과

6.1.1 군집분석 결과

군집분석을 4가지 변수(차원)를 기준으로 수행하였다. 구체적으로는 <Table 4>의 각 행을 분석 단위로 설정하였으며, 연도별 '사례 수(Cases filed)', '승인(Affirmative)된 사례 수', '반려(Negative)된 사례 수', 그리고 '중단(Terminated)된 사례 수' 네 가지 항목을 변수로 만들어 분석을 하였다. 사전에 하나의 변수가 비정상적으로 큰 영향력을 미치는 상황을 방지하기 위하여 각 변수를 조정(Feature scaling)하였다. 아래 식처럼 각 변수(x_i)에서 그 평균(μ_i)을 차감한 후 표준편차(σ_i)로 나누어 각 변수의 값들이 대략적으로 평균 0, 표준편차 1의 분포를 따르도록 설정하였다[47].

$$x_i^{'} = \frac{x_i - \mu_i}{\sigma_i}$$

군집분석 결과는 <Table 6>과 같다. 임의로 2개의 군집을 군집 1, 군집 2로 명명하였다. 총 28개 연도 중 21개 연도가 군집 1로 분류되었고, 7개 연도가 군집 2로 분류되었다. <Table 7>은 군집 별 주요 통계량을 정리해 놓은 것이다. 군집 1의 평균 사례 수, 평균 승인 횟수, 평균 수입 금액 등이 군집 2에 비해 상대적으로 높은 수치를 보이는 것으로 미루어 보아 군집 1의 연도들이 보호주의적인 경향이 강화되었던 기간을 의미한다고 유추해볼 수 있다.

〈Table 6〉 Clustering Results

Year	Cluster	Year	Cluster
1980	1	1994	2
1981	2	1995	2
1982	1	1996	2
1983	2	1997	2
1984	1	1998	2
1985	1	1999	2
1986	1	2000	2
1987	2	2001	1
1988	2	2002	2
1989	2	2003	2
1990	2	2004	2
1991	2	2005	2
1992	1	2006	2
1993	2	2007	2
		2008	2

〈Table 7〉 Major Statistics of Each Cluster

Cluster	Number of Observations	Average Number of Cases	Average Number of Affirmatives	Average Number of Negatives	Average Number of Terminations	Average Import Value
Cluster 1	7	116.57	36.71	48.85	35.28	1,329,050
Cluster 2	22	37.09	17.27	13.27	6.54	722,753

〈Table 8〉 Average Increase in Operating Incomes

Cluster		Cluster 1	Cluster 2	Difference (1-2)
Total	Gross Margin Growth	36.91	35.63	1.28
	Operating Income Growth	30.19	119.47	-89.28
Petroleum	Gross Margin Growth	16.27	32.57	-16.3
	Operating Income Growth	30.66	30.9	-0.24
Shipbuilding	Gross Margin Growth	107.47	44.21	63.26
	Operating Income Growth	56.82	123.78	-66.95
Automobiles	Gross Margin Growth	3.43	25.19	-48.76
	Operating Income Growth	22.51	154.28	-131.77
Automobiles Compoments	Gross Margin Growth	13.36	25.56	-12.20
	Operating Income Growth	10.77	168.92	-158.15

군집분석 후에, 군집별 평균 매출총이익 증가율과 평균 영업이익 증가율을 비교해 보았다. 우선 4개 산업 전체적인 증가율을 비교해 본 후, 석유정제, 조선, 자동차, 그리고 자동차 부품 산업으로 각각 나누어 군집 간 평균 이익 증가율을 비교해 보았다. 주요 결과는 <Table 8>과 같다.

이익 증가율 비교 결과 군집 1과 군집 2의 차이가 매출총이익 증가율에 비해 영업이익 증가율에서 두드러지게 나타나는 경향을 보여주었다. 4개 산업 영역을 통틀어 전체적으로 보았을 때, 매출총이익 증가율은 유의미한 차이를 보이지 않았지만(1.28%), 영업이익 증가율은 상당히 큰 차이를 보여주었다(-89.28%). 즉, 보호주의적인 무역 성향을 보이지 않은 군집 2의 기간이 미국의 보호주의적 무역 성향을 보인 군집 1의 기간에 비해 평균적으로 89.28% 나 높은 영업이익 증가율을 기록하였다.

이러한 경향은 조선, 자동차, 그리고 자동차 부품 산업에서 뚜렷하게 나타났다. 조선 산업은 군집 1에 비해 군집 2에서 66.95%, 자동차 산업은 131.77%, 자동차 부품 산업은 158.15% 높은 영업이익 증가율을 보여주었다. 하지만 조선 산업에서는 매출총이익 증가율은 오히려 군집 2에 비해 군집 1에서 높게 나타났다. 반면 자동차 및 자동차 부품 산업에서는 매출총이익 증가율도 군집 2에 비해 군집 1에서 높은 수치(각각 48.76%와 12.20% 차이)를 보여주었다.

결론적으로, 미국의 보호주의적 무역 행태에 가장 큰 영향을 받는 울산지역의 주력산업은 자동차와 관련된 산업이며, 이러한 영향을 잘 나타내는 지표는 영업이익 증가율이라고 할 수 있다. 미국 수출에 높은 의존도를 보이고

있는 현대 및 기아자동차는 최근 트럼프 미 대통령의 취임과 함께 강화되고 있는 보호무역주의 움직임에 대응하기 위해 미국 현지에 공장 시설을 확장하고 고용을 늘리는 등 투자를 급격히 확대하고 있다. 2017년 5월 현재 현대자동차와 기아자동차가 미국 앨라배마(Alabama) 및 조지아(Georgia) 공장 건설에 투자한 금액은 22억 달러(2조 5,000억 원)를 상회하며, 한국의 자동차 부품업체들이 미국 현지에서 80개가 넘는 공장을 운영하고 있는 것으로 알려져 있다. 이러한 경향을 반영하듯이, 지난 2016년에는 현대자동차와 기아자동차가 미국에서 판매한 자동차 중 수출 차량(714,000대)과 현재 생산 차량(709,000대)의 숫자가 엇비슷했다[13, 14, 35].

6.1.2 의사결정나무 분석 결과

군집분석에서 나아가, 미국 무역정책의 변화가 울산지역 주력산업에 미치는 영향을 예측하기 위하여 울산지역 주력산업별로 의사결정나무 분석(Decision Tree Analysis)을 수행하였다. 각 주력산업 별로 미국 보호주의 무역정책의 주요 지표인 연도별 '사례 수(Cases filed)', '승인(Affirmative)된 사례 수', '반려(Negative)된 사례 수', 그리고 '중단(Terminated)된 사례 수'를 통해 영업이익 증가율을 예측해 보고자 하였다. 분석의 결과물인 의사결정나무를 가지고, 미국의 특정 무역정책 기조 하에서 울산지역 주력산업의 영업이익 증가율, 즉 전반적인 성장률을 가늠해볼 수 있을 것으로 예상된다.

의사결정나무 분석의 과적합(overfitting)을 막기 위해 최대 깊이(depth)를 4로 한정하였다. 즉, 트리가 가장 깊게 자라도 4층을 넘어가지 않게 제한하였다. 아래 의사결정나무에서 나오는 'Cases'는 총 사례 수, 'Affirm'은 승인된 사례 수, 'Neg'는 반려된 사례 수, 그리고 'Term'은 중단된 사례 수를 의미하는 변수이다. 의사결정나무에서 각 노드의 맨 위 줄의 부등호가 노드의 *if-then* 분기조건이며, 분기조건이 만족되면(True) 좌측 자식 노드로 내려가고 분기조건이 만족되지 않으면(False) 우측 자식 노드로 내려가면서 진행하면 된다.

석유정제 산업의 영업이익 성장률을 예측하기 위한 의사결정나무의 모든 경우의 수에 따른 영업이익 성장률 예측치를 정리해 놓은 것이 <Table 9>이다. 석유정제 산업의 영업이익 성장률을 예측하기 위한 의사결정나무는 부록에 첨부하였다.

〈Table 9〉 Summary of Decision Tree to Predict the Business Performance of Petroleum Industry

Cases				Operating Income Growth (%)
Negative <= 1.5				**187.689**
Negative > 1.5	Number of cases <= 41.5	Number of cases <= 27.5	Terminated <= 8.0	**57.6089**
			Terminated > 8.0	**-1.2163**
		Number of cases > 27.5	Number of cases <= 30.5	**249.9683**
			Number of cases > 30.5	**53.3281**
	Number of cases > 41.5	Number of cases <= 118.0	Number of cases <= 46.0	**-40.1141**
			Number of cases > 46.0	**-11.5054**
		Number of cases > 118.0	Number of cases <= 126.5	**55.6311**
			Number of cases > 126.5	**32.5816**

〈Table 10〉 Summary of Decision Tree to Predict the Business Performance of Vessels Industry

Cases				Operating Income Growth(%)
Number of cases <= 13.0	Negative <= 47.0			1471.4746
	Negative > 47.0			353.599
Number of cases > 13.0	Negative <= 13.0	Affirmative <= 22.0	Negative <= 10.0	25.6203
			Negative > 10.0	247.8422
		Affirmative > 22.0	Number of cases <= 28.5	510.0909
			Number of cases > 28.5	268.928
	Negative > 13.0	Terminated <= 3.0		-340.9672
		Terminated > 3.0	Number of cases <= 34.0	65.9253
			Number of cases > 34.0	-17.0012

만약 1년 동안 반려된 사례의 수가 1.5개 미만이면 영업이익 성장률을 187.689%로 예측할 수 있다. 하지만 이러한 경우는 29년 동안 단 한 차례(1980)만 발생했으므로 예외적인 사례로 간주할 수 있다. 보다 일반적인 경우를 보자면, 연간 반려된 사례 수가 2회 이상이며, 연간 총 사례 수가 41개 이상 46개 이하면 영업이익 성장률을 -40.1141%로 예측해볼 수 있다. 이 경우에 해당하는 관측치(연도)의 수는 7개이며 오차의 평균제곱합은 591.096이다. 총 사례 수가 127개 이상으로 굉장히 많은 경우 영업이익 성장률은 32.5816%로 예측된다. 이 경우는 두 차례(1982년, 1992년) 발생하였으며 오차의 평균제곱합은 51.5585이다.

석유정제 산업의 의사결정나무를 보다 자세히 살펴보면, 대부분의 노드들이 총 사례 수(cases)에 의해 분기되는 것을 볼 수 있다. 실제로 <Table 9>의 모든 경우의 수를 보면 총 사례 수에 의해 영업이익 증가율이 매우 민감하게 반응하는 것을 볼 수 있다. 그러므로, 석유정제 산업의 성장률을 예측할 때에는 그 해에 발생한, 혹은 발생할, 총 반덤핑 관세 혹은 상계 관세 관련 사건의 수를 조사하는 것이 가장 유익할 수 있다.

<Table 10>은 조선 산업의 영업이익 성장률을 예측하기 위한 의사결정나무의 모든 경우의 수를 표현한 것이다. 예를 들어, 만약 총 사례 수가 13개 이상이고 승인된 사례 수가 22개 이하, 반려된 사례 수가 10개 이하면 해당 연도의 영업이익 성장률은 25.62%로 예측해 볼 수 있다. 해당 조건을 충족하는 연도는 9개 년도이며 오차의 평균제곱합은 3535.1763이다.

조선 산업의 의사결정나무를 조금 더 자세히 살펴보면, 석유정제 산업의 의사결정나무와는 다르게 총 사례 수, 반려된 사례 수, 승인된 사례 수 및 중단된 사례 수가 분기 조건에 골고루 분포되어 있는 것을 볼 수 있다. 그러므로, 조선 산업의 영업이익 성장률을 예측하고자 할 때에는 미국의 무역정책과 관련된 여러 가지 지표들을 보다 종합적으로 고려할 필요가 있다.

<Table 11>는 자동차 산업의 영업이익 성장률을 예측하기 위한 의사결정나무의 모든 경우의 수를 표현한 것이다. 자동차 산업의 영업이익 성장률을 예측하는 데 있어서는 도중에 중단된 사례(Term)의 수가 중요한 지표로 작용하는 것을 볼 수 있다. 예를 들어, 연간 중

〈Table 11〉 Summary of Decision Tree to Predict the Business Performance of Car Industry

Cases				Operating Income Growth(%)
Terminated <= 1.5	Terminated <= 0.5			76.8179
	Terminated > 0.5			2937.2534
Terminated > 1.5	Terminated <= 3.5	Number of cases <= 40.5	Number of cases <= 31.0	78.6795
			Number of cases > 31.0	31.6511
		Number of cases > 40.5	Number of cases <= 46.0	-171.3244
			Number of cases > 46.0	-32.8607
	Terminated > 3.5	Terminated <= 8.0	Terminated <= 6.0	30.3216
			Terminated > 6.0	184.0141
		Terminated > 8.0	Negative <= 15.0	27.3745
			Negative > 15.0	1.1654

〈Table 12〉 Summary of Decision Tree to Predict the Business Performance of Car Components Industry

Cases				Operating Income Growth(%)
Number of cases <= 57.5	Affirmative <= 51.0	Number of cases <= 43.0	Terminated <= 9.0	49.6541
			Terminated > 9.0	11.8731
		Number of cases > 43.0	Negative <= 10.5	-96.5624
			Negative > 10.5	9.2115
	Affirmative > 51.0	Affirmative <= 55.5		120.0909
		Affirmative > 55.5		6.9063
Number of cases > 57.5	Number of cases <= 68.5			3288.8528
	Number of cases > 68.5	Terminated <= 5.5	Negative <= 2.5	-13.8074
			Negative > 2.5	-2.3043
		Terminated > 5.5	Number of cases <= 118.0	24.5842
			Number of cases > 118.0	57.1445

단된 사례 수가 4 이상 6 이하면 자동차 산업의 영업이익 성장률을 30.3216%로 예측해 볼 수 있다.

<Table 12>은 자동차 부품 산업의 영업이익 성장률을 예측하기 위한 의사결정나무의 모든 경우의 수를 표현한 것이다. 자동차 부품 산업의 영업이익 성장률을 예측하기 위한 의사결정나무에는 총 사례 수, 승인된 사례 수, 반려된 사례 수, 그리고 중단된 사례 수 네 가지 지표가 가장 골고루 분기조건으로 들어간 의사결정나무라고 할 수 있다. 예를 들어, 총 사례 수가 43개 이하이고 승인된 사례 수가 51개 이하, 중단된 사례 수가 9개 이하이면 자동차부품 산업의 영업이익 성장률을 49.6541%로 예측해 볼 수 있다. 또 다른 예로, 총 사례 수가 69개 이상이고, 중단된 사례 수가 5개 이하, 반려된 사례 수가 2개 이하면 연간 영업이익 성장률을 -13.8074%로 예측해 볼 수 있다.

6.1.3 순환형 신경망 분석 결과

울산지역 주력산업의 매출총이익을 예측하기 위하여 과거 데이터를 활용해 순환형 신경망 모델의 일종인 장단기 기억 네트워크 모델을 학습하고 검증을 시도하였다. 29년간 4개 산업 분야의 총 96개의 데이터를 교차검증(cross validation)을 통해 각 경우의 수에 3번씩 학습하고 검증해 모델의 일반화 가능성을 테스트 해보았다. 매출총이익을 예측하기 위한 변수로는 이전 4년간 상계 관세 및 반덤핑 관세 관련 수치들과 매출총이익, 영업이익, 경상이익, 그리고 당기순이익 등 기업의 경영성과와 관련된 수치들을 활용하였다.

앞서 언급하였듯이, 검증 오차는 평균제곱근 편차(RMSE)로 계산하였으며 그리드 탐색(grid search) 방식을 통해 총 24가지 모델을 테스트해보았다. 즉, 하이퍼 파라미터의 각 경우의 수마다 하나씩 모델을 세워 전체 데이터를 3등분해 학습 데이터와 검증 데이터를 바꿔가며 학습과 검증을 반복해 검증 오차를 평균내어 계산한 것이다. 그리드 탐색 검증 결과는 <Table 13>과 같다. 검증 결과 셀의 크기가 10, 드랍아웃 비율이 0.2, 학습 에포크가 2000, 그리고 배치의 크기가 15일 때 가장 검증 오차가 1061.34로 가장 낮게 나타났다. 오차 계산 방식이 제곱합임을 감안해 보면 이는 상당히 높은 확률로 울산지역 산업의 성장 경향을 예측함 수 있음을 의미한다고 볼 수 있다.

즉, 우리는 장단기 기억 네트워크 모델을 통해 순차형 데이터를 효과적으로 모델링함으로써, 이전 4년간의 데이터를 활용해 미래의 울산지역 기업들의 매출총이익을 효과적으로 예측할 수 있다는 사실을 확인할 수 있었다. 향후 연구에서 보다 많은 연도의 데이터가 확보되고, 월 단위 혹은 분기 단위의 세밀한 데이터가 제공된다면 보다 높은 정확도를 기록할 수 있을 것으로 예상해 본다.

〈Table 13〉 Grid Search Results of LSTM Model

Size of cell	Dropout Rate	Epochs	Batch Size	Validation Error
10	0.2	1000	10	1291.60
10	0.2	1000	15	1426.96
10	0.2	1500	10	1335.56
10	0.2	1500	15	1351.53
10	0.2	2000	10	1832.48
10	0.2	2000	15	1061.34
10	0.4	1000	10	1954.37
10	0.4	1000	15	1756.83
10	0.4	1500	10	1243.54
10	0.4	1500	15	2030.88
10	0.4	2000	10	1502.58
10	0.4	2000	15	1963.74
20	0.2	1000	10	1560.41
20	0.2	1000	15	2141.41
20	0.2	1500	10	1776.82
20	0.2	1500	15	2417.95
20	0.2	2000	10	1470.76
20	0.2	2000	15	1602.83
20	0.4	1000	10	1759.57
20	0.4	1000	15	1953.03
20	0.4	1500	10	2187.49
20	0.4	1500	15	1730.04
20	0.4	2000	10	1502.58
20	0.4	2000	15	2659.24

6.2 분석 결과 요약

이상 미국 정부의 무역정책이 울산지역의 주력산업(석유정제, 조선, 자동차, 자동차부품)의 성장률에 미치는 영향을 살피기 위해 군집

분석과 의사결정나무 분석을 실시하였다. 군집 분석 결과, 미국의 보호주의적인 무역 성향이 다소 강하게 나타났던 군집(군집 1)과 그렇지 않았던 군집(군집 2)이 연간 영업이익 증가율에서 상당한 차이(89.28%)를 보여주었다. 이러한 경향은 조선(66.95%), 자동차(131.77%), 그리고 자동차 부품(158.15%) 산업에서 뚜렷하게 나타났으며, 특히 자동차와 자동차 부품 산업은 군집 간에 매출총이익 증가율에서도 상당한 차이를 보여주었다. 최근 미 트럼프 대통령 취임 및 미국의 보호주의로의 회귀와 함께 나타나고 있는 국내 자동차 및 자동차 부품 기업들의 미국 현지화 경향은 이러한 분석 결과와 맥을 같이한다고 할 수 있다.

나아가, 향후 미국 무역정책의 변화가 울산지역의 주력산업의 경영성과에 미치는 영향을 예측하기 위한 시도로 의사결정나무 분석과 순환형 신경망(장단기 기억 네트워크) 분석을 실시하였다. 미국 무역위원회의 반덤핑 및 상계 관세 사례와 관련된 통계 자료로 분석을 해본 결과, 각 산업별로 무역 지표를 통해 영업이익 증가율을 예측해 볼 수 있는 의사결정나무를 만들어 낼 수 있었다. 의사결정나무 분석 결과에서 유념할 점이 있다면 우리의 직관과는 다르게 각 산업별로 경영성과를 예측하는 데 유의미한 지표들이 다르다는 것이다. 결과를 해석하고 활용할 때 이러한 사실을 염두에 두어야만 오류에 빠지지 않고 객관적인 분석을 이어나갈 수 있다. 마지막으로, 장단기 기억 네트워크를 활용해 매출의 순차적 특성을 고려한 결과 좋은 예측 모델을 만들 수 있었다. 앞으로 보다 정확한 예측 모델을 만들기 위해서는 데이터의 이러한 특성을 고려해야 한다는 사실을 확인해 볼 수 있었다.

7. 결 론

7.1 요약

본 연구에서는 중화학 공업을 중심으로 울산지역의 주력산업과 미국의 보호주의 무역정책에 대해 개괄적으로 알아보고 둘 사이의 관계를 반덤핑 관세와 상계 관세 등 수입 상해 사건을 배경으로 탐구해 보았다.

우선 미국의 보호주의 무역정책 기조에 따른 울산지역의 기업 경영성과의 변화 양상을 알아보기 위해 군집분석을 활용하였다. k-평균 군집화 알고리즘을 활용하여 1980년부터 2008년까지 각 연도를 미국 무역 지표에 따른 두 개의 군집(보호주의 군집과 비보호주의 군집)으로 분류하여 둘 사이의 차이를 각 산업의 영업이익 증가율을 중심으로 살펴보았다. 그 결과 전반적으로 보호주의 성향이 강했던 연도의 영업이익 성장률이 그렇지 않았던 연도의 성장률에 비해 89% 가량 낮았던 것으로 밝혀졌다. 특히, 자동차와 관련된 기업이 이러한 차이를 보다 심하게 체감하는 것으로 나타났으며, 최근 자동차 관련 기업들의 미국 현지화 전략이 이러한 경향을 반영한다고 할 수 있다.

군집분석 후, 의사결정나무 분석을 통해 향후 미국의 무역정책 변화가 울산지역 주력산업의 경영성과에 미치는 영향을 예측하고자 하였다. 그 결과 직관적이고 이해하기 쉬운 결과를 도출할 수 있었으며, 앞으로 이러한 알고리즘을 활용하여 상계 관세와 반덤핑 관세를 중심으로 미국의 무역정책 변화에 따른 울산지역 기반 산업의 경영성과를 예측해 볼 수 있을 것으로 예상된다.

7.2 시사점 및 한계점

본 연구의 가장 큰 시사점은 미국 정부의 무역정책의 변화가 울산지역의 기업의 경영성과에 미치는 영향을 객관적으로 검증하고 어느 정도 정량화할 수 있다는 것이다. 우리나라와 미국 간의 밀접한 정치・외교・경제적 관계나 수출 의존적인 우리나라의 산업 특성을 미루어 보았을 때 대부분의 사람들이 직관적으로 미국의 무역정책과 수출 중심의 울산의 주력 산업의 경영성과가 상당 부분 관련되어 있을 것이라고 생각할 수 있다. 하지만, 본 연구에서는 이를 수치로 정량화하고 상관관계를 규명하기 위해 데이터에 기반한 기계학습 방법을 채택하였다. 그 결과, 미국의 반덤핑 및 상계관세 관련 지표가 우리나라 중공업 기업들의 회계 장부 상의 드러나는 장기적인 성과에 미치는 영향을 객관적인 숫자로 검증해볼 수 있었다.

기업과 정부에서는 본 연구의 분석 결과 및 향후 이에서 더욱 발전된 결과를 현실적인 계획 수립과 중장기적인 목표 달성에 충분히 활용할 수 있을 것으로 예상된다. 예를 들어, 미국의 보호주의적인 무역 성향이 강할 것으로 예상되는 시기에는 그로 인해 특별히 난항을 겪을 것으로 예상되는 산업을 보호하기 위한 정책을 펼치거나 특정 산업의 내수를 활성화시켜 외부효과로 인한 우리나라의 경제에 대한 전반적인 타격을 방지할 수 있다.

나아가, 이를 단순히 과거의 사건과 경향을 설명하는데 그친 것이 아니라 미래의경향을 가늠해볼 수 있는 예측 연구를 시도하였다. 본 연구의 결과를 보다 발전시켜 향후 연구에서는 축적된 데이터를 통해 미국 무역정책에 따라 달라지는 울산지역 주력산업의 경영성과를 보다 정확하게 예측하고 이를 현실에 반영할 수 있을 것으로 예상된다. 또한, 각 산업별로 경영성과에 미치는 요소들을 밝혀내는 과정 속에서 중요한 통찰을 얻을 수 있었다. 석유정제와 조선, 자동차, 그리고 자동차 부품 산업 모두 같은 무역 지표와 데이터를 사용하였지만, 각 산업별로 영업이익 성장률에 영향을 심대하게 미치는 요소는 다르게 나타났다. 그러므로, 데이터를 통해 기업의 경영성과 예측을 시도할 때에는 다양한 시각에서 종합적으로 접근을 해야 한다는 중요한 통찰을 확인해 볼 수 있었다.

많은 사람들이 이 세상에서 일어나는 일들에 대해 확신을 가지고 예측을 하지만 실제와 큰 괴리를 나타내는 경우가 적지 않다. 특히 2008년 금융위기와 같이 정치・경제・사회적으로 복잡한 문제의 경우 예측이 더욱 어려울 수 밖에 없다[48]. 극도로 불확실한 세상에서 예측의 정확도를 높이기 위한 하나의 방법으로 제안된 것이 ≪고슴도치(hedgehog)≫보다 ≪여우(fox)≫처럼 예측을 시도하는 것이다. 둘 사이에 상당히 많은 차이가 있지만 가장 큰 차이는 바로 여우는 고슴도치와는 다르게 열린 마음으로 다차원적인 사고를 한다는 것이다[30, 47, 48].

본 연구의 가장 큰 한계점은 학습에 활용된 데이터가 부족하다는 것이다. 일반적으로 기계학습 알고리즘이 현실에 바로 적용할 수 있는 성과를 보여주기 위해서는 이른바 ≪빅 데이터(Big data)≫가 마련되어야 하지만, 아직 미국 무역정책이나 울산지역의 주력 기업의 경영성과와 관련된 데이터는 가용할 수 있는 양이 많지 않고 체계화되어 정리되어 있지 못

하다. 앞으로 비정형 데이터(Unstructured data)를 포함한 다양한 데이터가 보다 많이 축적되고 가용 범위 내에 들어온다면 보다 믿을 수 있고 강건한 예측이 가능할 수 있을 것으로 보인다.

References

[1] Alsabti, K., Ranka, S., and Singh, V., An efficient k-means clustering algorithm, 1997.

[2] Asia Economy, http://view.asiae.co.kr/news/view.htm?idxno=2017051611153450902. 2017-05-27 accessed, 2017.

[3] Baek, C., Studies in US Trade Policy. Sapyoung Academy, 2015.

[4] Bank of Korea Financial Statement Analysis for 2015.

[5] Bank of Korea, Financial Statement Anaylsis for 2017.

[6] Korea International Trade Association, Export/Import Statistics, 2017.

[7] BBC News, Why is the US turning to protectionism?, http://www.bbc.com/news/world-us-canada-36368759. 2017-05-27 accessed, 2016.

[8] Business Insider, The country that imposes the most restrictions on trade might surprise you, http://www.businessinsider.com/the-us-is-the-most-protectionist-nation-2015-9?IR=T, 2017-05-27 accessed, 2015.

[9] Choi, J., Managing by Numbers, One and One Books, 2012.

[10] CNBC, Trump's protectionism biggest threat to US economy's growth, CNBC Fed Survey respondents say, http://www.cnbc.com/2017/01/30/trumps-protectionism-biggest-threat-to-us-economys-growth-cnbc-fed-survey-respondents-say.html, 2017-05-27 accessed, 2017.

[11] Colah. Understanding LSTM Networks, http://colah.github.io/posts/2015-08-Understanding-LSTMs, 2017-07-15 accessed, 2015.

[12] Destler, I. M., American trade politics. Columbia University Press, 2005.

[13] EBN, http://www.ebn.co.kr/news/view/879646, 2017-06-19 accessed, 2017.

[14] ET News, http://www.etnews.com/20170301000021, 2017-06-19 accessed, 2017.

[15] Ferguson, T., "From Normalcy to New Deal: industrial structure, party competition, and American public policy in the Great Depression," International Organization, Vol. 38, No. 1, pp. 41-94, 1984.

[16] Flach, P., Machine learning: the art and science of algorithms that make sense of data, Cambridge University Press, 2012.

[17] Frieden, J., "Sectoral conflict and foreign economic policy, 1914~1940," International Organization, Vol. 42, No. 1, pp. 59-90, 1988.

[18] Gokhale, S. S. and Lyu, M. R., "Regression tree modeling for the prediction of software quality," In: proceedings of the

Third ISSAT International Conference on Reliability and Quality in Design, pp. 31-36, 1997.

[19] Hartz, L., The Liberal Tradition in America, Houghton Mifflin Harcourt, 1991.

[20] Hochreiter, S. and Schmidhuber, J., "Long short-term memory," Neural computation, Vol. 9, No. 8, pp. 1735-1780, 1997.

[21] Jain, A. K., "Data clustering: 50 years beyond K-means," Pattern recognition letters, Vol. 31, No. 8, pp. 651-666, 2010.

[22] Junggi, http://www.junggi.co.kr/article/articleView.html?no=18925, 2017-05-27 accessed, 2017.

[23] Kanungo et al., "An efficient k-means clustering algorithm: Analysis and implementation," IEEE transactions on pattern analysis and machine intelligence, Vol. 24, No. 7, pp. 881-892, 2002.

[24] Karpathy, A. and Johnson, J., and FEI-FEI, L., "Visualizing and understanding recurrent networks," arXiv preprint arXiv: 1506.02078, 2015.

[25] Kbiz, http://news.kbiz.or.kr/news/articleView.html?idxno=43685 2017-05-27 accessed, 2017.

[26] Kim, H. and Kim, J., "A Study on Consumer Value Perception through Social Big Data Analysis: Focus on Smartphone Brands," The Journal of Society for e-Business Studies, Vol. 22, No. 1, pp. 123-146, 2017.

[27] Kriegel, H.-P., Schubert, E., and Zimek, A., "The (black) art of runtime evaluation: Are we comparing algorithms or implementations?," Knowledge and Information Systems, pp. 1-38, 2016.

[28] Kyeongsang Daily, http://www.ksilbo.co.kr/news/articleView.html?idxno=565631, 2017-06-22 accessed, 2016.

[29] Kyeongsang Daily, http://www.ksilbo.co.kr/news/articleView.html?idxno=591017, 2017-05-25 accessed, 2017.

[30] Lake, D. A., International Economic Structures and American Foreign Economic Policy, 1887~1934, World Politics, Vol. 35, No. 4, pp. 517-543, 1983.

[31] Lee, D., Won, J. W., Kwon, Y., and Kim, M., "A Study on Customer Satisfaction for Courier Companies based on SNS Big data," The Journal of Society for e-Business Studies, Vol. 21, No. 4, pp. 55-56, 2017.

[32] Lerner, J. S. and Tetlock, P. E., "Accounting for the effects of accountability," Psychological bulletin, Vol. 125, No. 2, p. 255, 1999.

[33] Loh, W. Y., Classification and regression tree methods. Encyclopedia of statistics in quality and reliability, 2008.

[34] Michaelsen et al., "Regression tree analysis of satellite and terrain data to guide vegetation sampling and surveys," Journal of Vegetation Science, Vol. 5, No. 5, pp. 673-686, 1994.

[35] Milner, H. V. and Yoffie, D. B., "Between free trade and protectionism: strategic trade policy and a theory of corporate

trade demands," International Organization, Vol. 43, No. 2, pp. 239-272, 1989.

[36] Mirkes, E. M., K-means and K-medoids applet, University of Leicester, http://www.math.le.ac.uk/people/ag153/homepage/KmeansKmedoids/Kmeans_Kmedoids.html. 2017-06-17 accessed, 2011.

[37] Naeil, http://www.naeil.com/news_view/?id_art=236233, 2017-06-19 accessed, 2017.

[38] Naver Terms, http://terms.naver.com/entry.nhn?docId=2085598&cid=48591&categoryId=48591, 2017-05-27 accessed, 2013.

[39] News One, http://news1.kr/articles/?2931643, 2017-05-25 accessed, 2017.

[40] New York Times, Why Breixt Is Best for Britain: The Left-Wing Case, https://www.nytimes.com/2017/03/28/opinion/why-brexit-is-best-for-britain-the-left-wing-case.html, 2017-05-25 accessed, 2017.

[41] Provost, F. and Fawcett, T., "Data science and its relationship to big data and data-driven decision making," Big Data, Vol. 1, No. 1, pp. 51-59, 2013.

[42] Safavian, S. R. and Landgrebe, D., "A survey of decision tree classifier methodology," IEEE transactions on systems, man, and cybernetics, Vol. 21, No. 3, pp. 660-674, 1991.

[43] Schattschneider, E. E., Politics, Pressures and the Tariff: A study of free private enterprise in pressure politics, as shown in the 1929~1930 revision of the tariff. Prentice-Hall, inc, 1935.

[44] Seoul Economy, http://www.sedaily.com/NewsView/1OARZOM3J8, 2017-05-25 accessed, 2017.

[45] Seoul Economy, http://www.sedaily.com/NewsView/1ODFK34YLB, 2017-05-27 accessed, 2017.

[46] Shaffer, C. A., "Data structures and algorithm analysis," 2012.

[47] Shmueli, G., Patel, N. R., and Bruce, P. C., "Data Mining for Business Analytics: Concepts, Techniques, and Applications with XLMiner," John Wiley & Sons, 2016.

[48] Silver, N., The signal and the noise: Why so many predictions fail-but some don't, Penguin, 2012.

[49] Taleb, N. N., The black swan: The impact of the highly improbable(Vol. 2), Random house, 2007,

[50] Tetlock, P., Expert political judgment: How good is it? How can we know?, Princeton University Press, 2005.

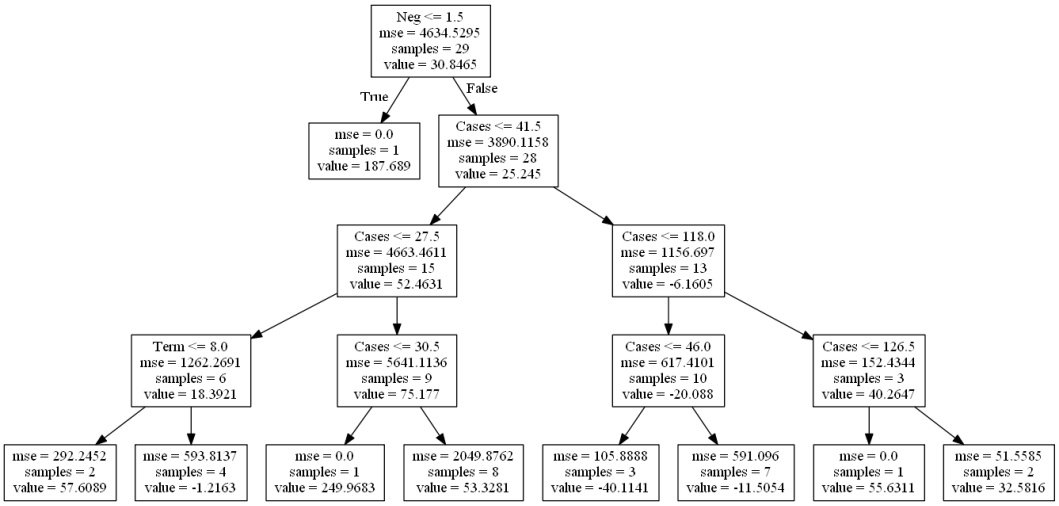
[51] Tutorialspoint, Data Structure and Algorithms-Tree, https://www.tutorialspoint.com/data_structures_algorithms/tree_data_structure.htm, 2017-06-19 accessed, 2017.

[52] Tutorialspoint, Data Mining-Decision Tree Induction, https://www.tutorialspoint.com/data_mining/dm_dti.htm, 2017-06-17 accessed, 2017.

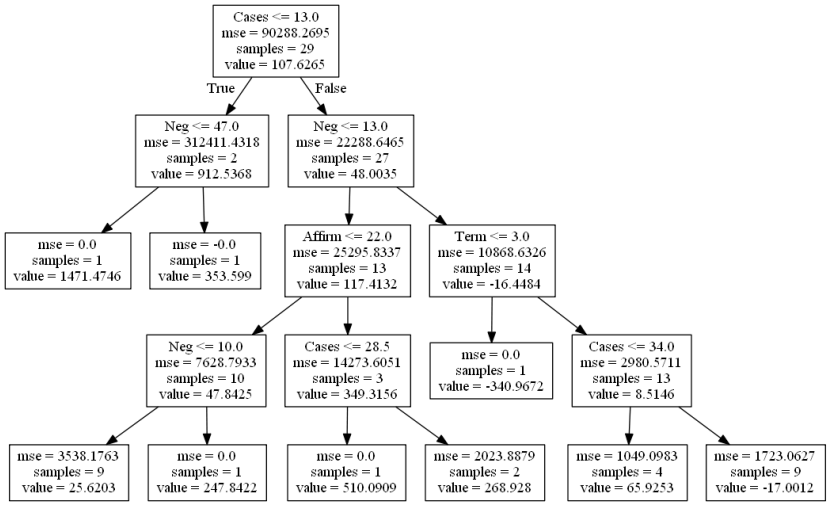
[53] Ulsan Daily, http://www.iusm.co.kr/news/articleView.html?idxno=738873, 2017-05-25 accessed, 2017.

[54] Ulsan Daily, http://www.iusm.co.kr/news/articleView.html?idxno=723597, 2017-05-25 accessed, 2017.
[55] Ulsan Jonghap Daily, http://www.ujnews.co.kr/news/articleView.html?idxno=263136, 2017-05-25 accessed, 2017.
[56] Unites States International Trade Commission, Antidumping and Countervailing Duty Investigations, https://www.usitc.gov/trade_remedy/731_ad_701_cvd/investigations.htm, 2017-05-27 accessed, 2017.
[56] Unites States International Trade Commission, Import Injury Investigations Case Statistics(FY 1980~2008), 2010.

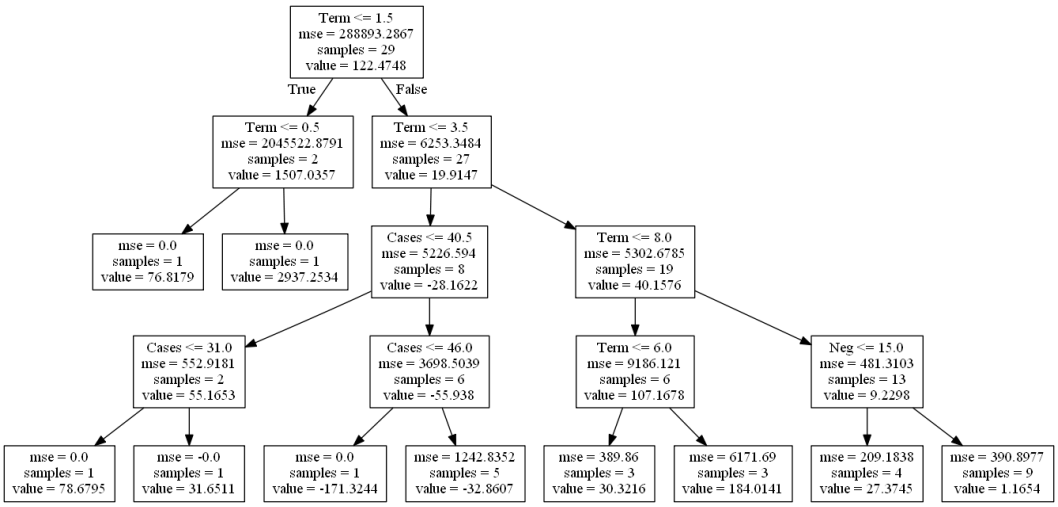
〈부 록〉



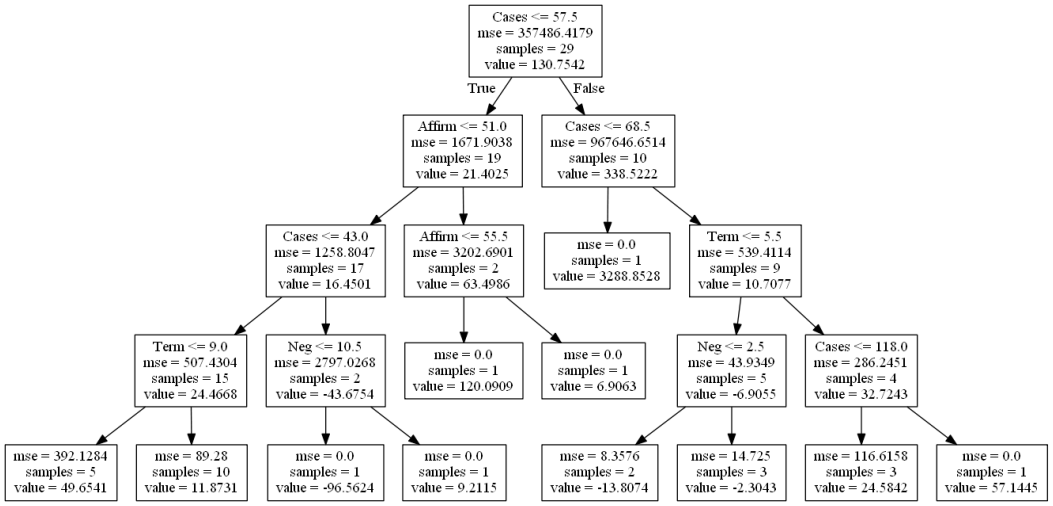
〈Decision Tree to Predict the Operating Income of Petroleum Industry〉



〈Decision Tree to Predict the Operating Income of Shipbuilding Industry〉



〈Decision Tree to Predict the Operating Income of Automobile Industry〉



〈Decision Tree to Predict the Operating Income of Automobile Components Industry〉

저 자 소 개

박진수 (E-mail: jinsoo@snu.ac.kr)
1999년 The University of Arizona (경영학 박사)
1999~2002년 Minnesota University 조교수
2002~2005년 고려대학교 조교수
2005년~현재 서울대학교 경영대학원 교수
관심분야 온톨로지, 빅데이터 분석, 데이터 통합 등

김경호 (E-mail: kimkyoungho@bok.or.kr)
2015년 서울대학교 경제학부 (경제학 학사)
2015년~현재 한국은행 조사역, 통화정책국, 울산본부
관심분야 국제경제학, 계량경제학, 화폐금융론 등

김범수 (E-mail: buomsookim0927@gmail.com)
2016년 서울대학교 경영대학 (경영학 학사)
2016년~현재 서울대학교 경영대학 (경영학 석사과정)
관심분야 데이터 마이닝, 인공지능

서지혜 (E-mail: jihaesuh77@gmail.com)
2000년 경희대학교 (영문학 학사)
2002년 경희대학교 경영대학 (경영학 석사)
2006년 Carnegie Mellon University (경영학 석사)
2017년 서울대학교 경영대학 (경영학 박사)
2017년~현재 도시 데이터 사이언스 연구소 연구원
관심분야 데이터 모델링, 데이터 마이닝