



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

경영학석사 학위논문

조선기자재 분야 소상공인 정책자금의
부실에 관한 연구

Study on Insolvency of Government Policy Loans for
the Small Enterprises in Marine Equipment Industries

지도교수 이재민

2019년 8월

한국해양대학교 해양금융·물류대학원

해양금융학과

최 미 나

본 논문을 최미나의 경영학석사 학위논문으로 인준함.

위원장 이 기 환 (인)

위 원 이 재 민 (인)

위 원 임 상 섭 (인)

2019년 7월

한국해양대학교 대학원

목 차

표 차례	iii
그림 차례	iv
Abstract	v
제 1 장 서 론	
1.1 연구의 배경 및 목적	1
1.2 연구의 방법과 구성	3
제 2 장 소상공인 현황	
2.1 소상공인의 개념	5
2.2 소상공인의 현황	7
2.3 소상공인 자금지원의 필요성	10
제 3 장 선행연구	
3.1 기업부실의 개념	12
3.2 부실예측 선행연구	14
제 4 장 연구방법	
4.1 랜덤 포레스트(Random Forest)	19
4.2 로지스틱 회귀분석(Logistic Regression)	22
4.3 평가지표	25
제 5 장 소상공인 정책자금 부실예측	
5.1 표본기업의 선정	27
5.2 변수 선정	28
5.3 실증분석	29
5.3.1 Random Forest 모형의 평가	30
5.3.2 Logistic Regression 모형의 평가	31

제 6 장 소상공인 정책자금 부실관리	
6.1 부실관리의 목적	32
6.2 변수선정	33
6.3 실증분석	33
6.3.1 Random Forest 모형의 평가	34
6.3.2 Logistic Regression 모형의 평가	35
6.4 부실관리 방안	35
 제 7 장 소상공인 정책자금 지원제도 개선 방안	 37
 제 8 장 결 론	 44
 참고문헌	 46

표 차례

[표 1] 상시근로자 수의 산정	5
[표 2] 소상공인 기준	6
[표 3] 소상공인 사업체수와 비중	7
[표 4] OECD회원국의 자영업자 비율 (2011년~2017년)	9
[표 5] 1인 창조기업 경영 애로사항 1순위	10
[표 6] 소상공인 창업 애로사항	11
[표 7] 소상공인 경영 애로사항	11
[표 8] 재무적 부실징후	13
[표 9] 다변량판별분석, 로짓모형을 이용한 선행연구	15
[표 10] 머신러닝을 이용한 선행연구	17
[표 11] Confusion Matrix	25
[표 12] 분석에 사용된 변수	28
[표 13] Random Forest 모형의 예측 결과	30
[표 14] Logistic Regression 모형의 예측 결과	31
[표 15] 소상공인 생존율(평균)	32
[표 16] Random Forest 모형의 예측 결과	34
[표 17] Logistic Regression 모형의 예측 결과	35
[표 18] 사고기업 분포비율	38
[표 19] 부실기업의 2금융권 대출 이용 현황	39
[표 20] 대부업·사금융시장이용자 설문조사	42

그림 차례

[그림 1] 기업규모별 종사자 수	7
[그림 2] 연도별 소상공인 현황	8
[그림 3] OECD회원국의 자영업자 비율(2018년)	9
[그림 4] 랜덤포레스트(Random Forest)의 구조	21

Study on Insolvency of Government Policy Loans for the Small Enterprises in Marine Equipment Industries

Choe, Mi Na

Department of Marine Finance

Graduate School of Korea Maritime and Ocean University

Abstract

Insolvency of enterprises not only causes serious damages to the management of enterprises as well as the stakeholders such as investors, subcontractors or clients but also incurs significant losses to overall national economy. Since the number of small businesses, in particular, accounted for a significantly high portion, over 85% of the entire enterprises across the country, it is necessary to predict any potential crisis of enterprises in advance and prepare for that.

Even if studies on prediction for insolvency of enterprises have been performed for long time in both home and abroad, the empirical studies on insolvency of enterprises have not been performed actively due to either a lack of interests or hardships in relevant data collection. However, the small enterprises are the critical economic entity in which creates employment, activates local economy and contributes to a development of national economy. In addition, the government has supported the business categories that are

sensitive on economic tendency during the period of economic recession with much more budget, accordingly, in these contexts, a study on the small enterprises that are playing in the marine equipment industries, which is one of government supported categories, is absolutely required. Therefore, this study aims to empirically analyze the factors that affect an insolvency of small enterprises in order for making a contribution to an improvement in the credit evaluation and insolvency management institutions of government policy loans for the small enterprises.

For empirical analysis, information about the enterprises that received a government policy loan for the small enterprises were collected from the Agency for Traditional Market Administration in order to constitute samples. Afterwards, collected data were analyzed using a logistic regression analysis and a random forest that is one of machine learning techniques.

In the results of this study, a high prediction result was observed in a logistic regression model. In other words, the model conducted an experiment with combination of financial and non-financial factors exhibited a higher prediction probability than the models that conducted experiments with financial factors or non-financial factors alone. Based on that, it was possible to confirm a need for development of non-financial factors. On the other hand, it is suggested that the model with poor prediction results needs to collect more detail information about the small enterprises and refine the credit evaluation model more elaborately.

KEY WORDS : small enterprise; Government policy loans; prediction for insolvency; logistic regression; random forest

조선기자재 분야 소상공인 정책자금의 부실에 관한 연구

최 미 나

한국해양대학교 해양금융·물류대학원

해양금융학과

요 약

기업의 부실은 기업의 경영자뿐만 아니라 투자자, 거래처 등 기업의 이해관계자들에게 손실을 끼치고 국가 경제에 전반적으로 피해를 주게 된다. 특히 소상공인은 그 수가 전체사업체의 85%이상에 해당하는 높은 비중을 차지하고 있으므로 기업의 위기를 사전에 예측하고 대응할 필요성이 있다.

기업 부실예측에 대한 연구는 국내외에서 활발히 진행되어 왔으나 소상공인의 부실에 대한 실증적 연구는 관심의 부재나 자료수집의 어려움 등으로 연구가 활발하지 못하였다. 하지만 소상공인은 고용을 창출하고 지역경제를 활성화며 국가경제의 발전에 기여하는 중요한 경제의 주체이다. 게다가 경기 불황시기에 경기민감 업종을 대상으로 더 많은 정부예산이 투입되어왔으므로 그 중 하나인 조선기자재 분야 소상공인에 대한 연구의 필요성이 크다. 그러므로 본 연구에서는 기업 부실에 영향을 미치는 요인을 실증 분석하여 소상공인 정책자금의 여신심사와 부실관리 제도의 개선에 기여하고자 한다.

본 연구는 소상공인시장진흥공단에서 소상공인 정책자금을 지원받은 기업의 정보를 수집하여 표본을 구성하였으며 머신러닝 기법중 하나인 랜덤포레스트와 로지스틱회귀분석을 이용하였다. 연구결과 로지스틱회귀모형에서 높은 예측 결과를 얻을 수 있었는데, 재무적 변수로만 실험한 모형보다 재무적 변수와 비재무적 변수를 혼합하여 실험한 모형이 더 높은 예측확률을 보임으로써 비재무적 지표 개발의 중요성을 확인할 수 있었다. 반면 예측결과가 좋지 않았던 모형은 기업정보가 더 자세히 수집되고 여신심사평가모형이 더욱 정교화 될 필요성이 있음을 시사한다.

키워드 : 소상공인; 정책자금; 부실 예측; 로지스틱회귀분석; 랜덤포레스트

제 1 장 서론

1.1 연구의 배경 및 목적

소상공인은 경제의 뿌리 또는 핏줄이라는 말로 비유되곤 한다. 소상공인은 뿌리와 핏줄처럼 지역 구석구석에 포진되어 고용을 창출하고 지역 경제를 활성화하며 국가경제의 발전에 기여하는 중요한 경제의 주체이다. 2016년 조사자료에 의하면 소상공인은 우리나라 전체 사업체 수의 85.3%, 전체 종사자 수의 36.4%를 차지하고 있다. 그러나 그 중요성에 비하여 상대적으로 정부의 지원에 있어 소외된 측면이 있다. 소상공인은 대기업과 비교하여 경영여건, 재무구조, 자금조달 여력 등 여러 측면에서 열악하다. 그러므로 내수부진, 경기침체 등 외부환경이 나빠질 경우 매출감소, 영업이익률 하락 등의 과정을 거쳐 폐업에 이르는 사업체가 많다. 소상공인은 규모가 큰 대기업, 중견기업에 비하여 생존률이 상당히 낮다. 기업의 부실은 금융기관, 거래처, 종사자 등 기업의 이해관계자에게 영향을 주게 되는데, 특히 소상공인은 사업체의 수가 많고 폐업률도 높기 때문에 막대한 사회적 비용을 초래할 수 있으므로 부실을 예측하여 피해를 최소화할 필요성이 있다.

본 연구의 대상은 조선기자재 분야 제조업을 영위하는 소상공인이다. 우리나라의 급속한 경제발전과정에서 조선업은 일자리의 창출과 수출에 기여하며 중요한 역할을 담당해왔고 조선기자재 산업도 마찬가지로 역할을 하였다. 조선기자재 산업은 조선업의 후방산업이며 조선업은 경기민감업종에 해당하여 경기의 영향을 많이 받는다. 조선기자재 산업은 선박의 건조에 사용되는 기계류, 자재류, 선박을 구성하는 부분품 등을 포함하는 산업이다. 따라서 경기변동에 의해 조선업이 침체되면 조선기자재

산업도 침체된다. 그동안 경기 불황시기에는 정부가 경기민감업종 및 협력기업을 대상으로 하는 특별자금을 지원하여 조선기자재 관련 산업에는 다른 산업군보다 많은 정부예산이 투입되어왔다. 그러므로 심도 있는 연구가 이루어져야함에도 조선기자재 산업에 초점을 맞춘 기업부실에 관한 선행연구는 전무한 상황이다.

소상공인의 재무적 정보는 대기업처럼 전담인력에 의해 체계적으로 관리되지 못하고 신뢰성이 낮아 여신심사와 부실관리에 활용하기에 부족함이 많다. 그러므로 민간 금융기관에서는 신용을 평가할 수 있는 정보가 부족하여 신용대출보다는 주로 담보대출이 취급되었다. 그러나 소상공인 정책자금 지원정책이 일회성으로 끝나는 것이 아니므로 이러한 정보부족의 문제점들이 근본적으로 개선될 필요성이 있다. 창업에 의한 고용창출과 소상공인의 경영안정화를 위한 정부의 정책에 따라 소상공인 정책자금 지원규모는 매년 증가하고 있고, 지원규모 확대에 따라 부실규모도 증가추세에 있으므로 부족한 재무적 정보를 보완할 수 있는 대책마련이 시급하다.

경기가 호황일 때는 정책금융기관뿐만 아니라 민간 금융기관에서의 자금조달도 어렵지 않다. 그러나 경기가 불황일 때 규모가 작은 기업은 자금을 조달하는 것이 상당히 어렵다. 경기가 불황일수록 다른 기관에서 자금을 조달하기 어렵기 때문에 정부는 불황일 때 더 적극적으로 자금을 지원해야한다. 정책자금의 지원목적은 고려하면 경기역행적으로 운용되어야 하지만 경제 상황이 좋지 않으면 기업의 부실 가능성이 높아질 수밖에 없으므로 보수적으로 심사하는 경향이 있다. 두 마리 토끼를 다 잡는 것은 어렵지만 정책자금의 손실을 최소화하고 자금지원의 효과를 높이기 위해서는 심사평가모형이 지금보다 더 정교화되어야 할 것이다. 심

사모형의 정교화는 선별하여 기회를 제한하는 목적보다는 기회를 찾기 위한 목적이 더 크다. 대출 신청기업의 재무적 상황이 좋지 않지만 그럼에도 불구하고 어떤 정보에 의해 자금지원이 가능한지, 어떤 점이 우수하므로 발전가능성이 높은지 등을 발견하기 위한 목적이다. 전통적인 평가방식에 의해 자금을 지원받지 못한 기업이 일시적인 유동성부족으로 흑자도산이 되는 사례를 보면 평가방식의 개선이 필요함을 알 수 있다.

본 연구의 목적은 다음과 같다. 첫째, 재무적 정보와 비재무적정보를 이용하여 기업의 부실을 예측함으로써 여신심사와 부실 관리 제도를 보완하고자 한다. 비재무적 정보를 이용하여 실증연구를 시도하는 과정은 어떤 정보가 재무적 정보의 대안이 될 수 있는지에 대한 아이디어를 제공해 줄 수 있을 것이다. 둘째, 부실을 예방하고 부실로 인한 사회적 비용을 최소화하며 지속가능한 지원을 위해 소상공인 정책자금 지원제도 개선방안을 모색하고자 한다. 한정된 재원이 효율적으로 배분되어 정책의 효과를 높이기 위해서는 기존의 정책자금 지원기업의 사례를 검토하여 문제점을 개선하고 제도를 보완하는 과정이 필요하다. 소상공인 지원 개선 방안을 통해 대출소외계층의 금융지원이 활성화 되도록 하며 소상공인이 지속적으로 사업을 영위하는데 도움이 되고자 한다.

1.2 연구의 방법과 구성

선행연구는 주로 규모가 큰 기업을 대상으로 하며 재무적 요인에 초점을 둔 부실예측연구가 많았다. 그러나 본 연구는 소상공인의 특징을 반영하여 재무적 정보와 비재무적 정보를 활용한 연구를 진행하고자 한다. 규모가 작은 기업은 외부감사대상기업에 해당하지 않으므로 재무정보의 신뢰성이 낮은 편이다. 비외감 중소기업인 198,164개를 실증분석한 허광복

(2018)의 연구에서도 비외감 기업이 금융기관에서 자금을 조달할 때 세무서장 증명 재무제표 이외의 재무제표를 제출하는 경우 발생액 이익조정을 이용해 이익을 상향조정하는 것으로 나타났으며 비외감기업 정보의 신뢰성에 문제가 있음을 알 수 있다. 본 연구에서는 재무제표의 신뢰성 문제를 보완하기 위해 비재무적 정보를 함께 수집하여 실증분석을 진행하였다. 연구의 표본은 조선기자재 분야 제조업을 영위하는 소상공인이며 그동안 연구가 활발하지 못했던 경기민감업종 소상공인의 부실에 대한 실증연구를 수행하였다. 연구방법으로는 랜덤포레스트와 로지스틱 회귀분석을 사용하였다.

본 논문의 구성은 다음과 같다. 제1장은 연구의 배경과 목적, 연구의 방법과 구성에 대하여 기술하였으며, 제2장에서는 소상공인의 개념과 현황, 그리고 소상공인 자금지원의 필요성에 대하여 살펴보았다. 제3장에서는 기업부실의 개념과 부실예측에 대한 선행연구를 다루었으며 제4장은 연구방법론을, 제5장에서는 소상공인 정책자금 부실예측에 대한 연구 내용을 정리하였다. 제6장에서는 소상공인 정책자금 부실관리에 대한 연구 내용을 정리하였으며 제7장에서는 소상공인 정책자금지원제도 개선방안을 기술하였고 제8장에서는 연구결과와 연구의 한계 및 시사점을 제시하였다.

제 2 장 소상공인 현황

2.1 소상공인의 개념

『소상공인 보호 및 지원에 관한 법률』에 의하면 소상공인이란 「중소기업기본법」 제2조 제2항에 따른 소기업 중 업종별 상시 근로자 수가 대통령령으로 정하는 기준(주된 사업에 종사하는 상시근로자의 수가 광업, 제조업, 건설업, 운수업의 경우에는 10인 미만, 그 밖의 업종의 경우 5인 미만)에 해당하는 자이다. 상시 근로자는 근로자¹⁾ 중에서 아래에 해당하는 자를 제외한 자이다.

- ① 임원 및 일용근로자
- ② 3개월 이내의 기간을 정하여 근로하는 자
- ③ 기업부설연구소 및 연구개발전담부서의 연구전담요원
- ④ 단시간 근로자로서 1개월 동안의 소정 근로시간이 60시간 미만인 자

상시 근로자 수의 산정 방법은 크게 직전 사업연도의 사업기간이 12개월 이상인 기업과 12개월 미만인 기업으로 구분된다.

[표 1] 상시근로자 수의 산정

구 분		상시 근로자 수 산정 방법
직전사업연도의 사업기간이 12개월 이상인 기업		직전 사업연도의 매월 말일 기준 상시 근로자 수를 합하여 12로 나눈 인원
직전사업연도의 사업 기간이 12개월 미만인 기업	업력 12개월 이상	최근 12개월 동안의 매월 말일 기준 상시근로자수를 합하여 12로 나눈 인원
	업력 12개월 미만	창업일부터 산정일까지 기간의 매월 말일 기준 상시근로자수를 합하여 해당 월수로 나눈 인원
	업력 1개월 미만	산정일 현재의 상시근로자 수

기업의 업종이 아래 표의 최근 3년간 평균매출액 기준을 충족하면 소기업에 해당하고, 상시근로자 기준까지 충족하면 소상공인에 해당하여 소상공인 지원정책의 대상이 될 수 있다.

1) 직업의 종류와 관계없이 임금을 목적으로 사업이나 사업장에 근로를 제공하는 자(「근로기준법」 제2조제1항제1호)

[표 2] 소상공인 기준

해당 기업의 주된 업종	분류 기호	소기업 (최근3년간 평균매출액)	소상공인
식료품 제조업	C10	120억원 이하	상시근로자 10인 미만
음료 제조업	C11		
의복, 의복액세서리 및 모피제품 제조업	C14		
가죽, 가방 및 신발 제조업	C15		
코크스, 연탄 및 석유정제품 제조업	C19		
화학물질 및 화학제품 제조업(의약품 제조업은 제외한다)	C20		
의료용 물질 및 의약품 제조업	C21		
비금속 광물제품 제조업	C23		
1차 금속 제조업	C24		
금속가공제품 제조업(기계 및 가구 제조업은 제외한다)	C25		
전자부품, 컴퓨터, 영상, 음향 및 통신장비 제조업	C26		
전기장비 제조업	C28		
그 밖의 기계 및 장비 제조업	C29		
자동차 및 트레일러 제조업	C30		
가구 제조업	C32	80억원 이하	
담배 제조업	C12		
섬유제품 제조업(의복 제조업은 제외한다)	C13		
목재 및 나무제품 제조업(가구 제조업은 제외한다)	C16		
펄프, 종이 및 종이제품 제조업	C17		
인쇄 및 기록매체 복제업	C18		
고무제품 및 플라스틱제품 제조업	C22		
의료, 정밀, 광학기기 및 시계 제조업	C27		
그 밖의 운송장비 제조업	C31		
그 밖의 제품 제조업	C33		
광업	B		
건설업	F		
운수업	H		
농업, 임업 및 어업	A		
금융 및 보험업	K		
전기, 가스, 증기 및 수도사업	D	120억원 이하	
도매 및 소매업	G	50억원 이하	
출판, 영상, 방송통신 및 정보서비스업	J		
하수·폐기물 처리, 원료재생 및 환경복원업	E	30억원 이하	
부동산업 및 임대업	L		
전문·과학 및 기술 서비스업	M		
사업시설관리 및 사업지원 서비스업	N		
예술, 스포츠 및 여가 관련 서비스업	R	10억원 이하	
숙박 및 음식점업	I		
교육 서비스업	P		
보건업 및 사회복지 서비스업	Q		
수리(修理) 및 기타 개인 서비스업	S		

자료 : 중소기업기본법 시행령 별표 3 (개정 2017.10.17.)

2.2 소상공인의 현황

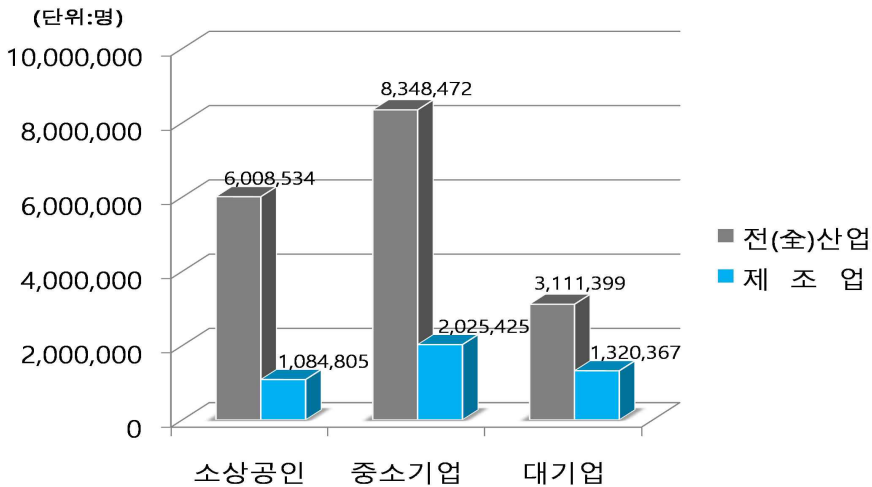
2016년 조사 자료를 살펴보면, 아래 표에서 소상공인의 사업체 수는 3,137,700개로 우리나라 전체 사업체 수의 85.3%를 차지하고 있으며 그 중 제조업은 소상공인의 사업체 수의 비중이 83.7%에 해당할 만큼 큰 비중을 차지하고 있다. 본 연구의 대상기업이 모두 제조업이므로 전체 업종과 제조업을 규모별로 비교하였다.

[표 3] 소상공인 사업체수와 비중

구분	전체 사업체수(개)	비중(%)	제조업 사업체수(개)	비중(%)
소상공인	3,137,700	85.3	345,825	83.7
중소기업	534,627	14.5	65,611	15.9
대기업	4,172	0.1	1,564	0.4
합계	3,676,499	100.0	413,000	100.0

자료 : 중소벤처기업부 통계자료(2016) 참고

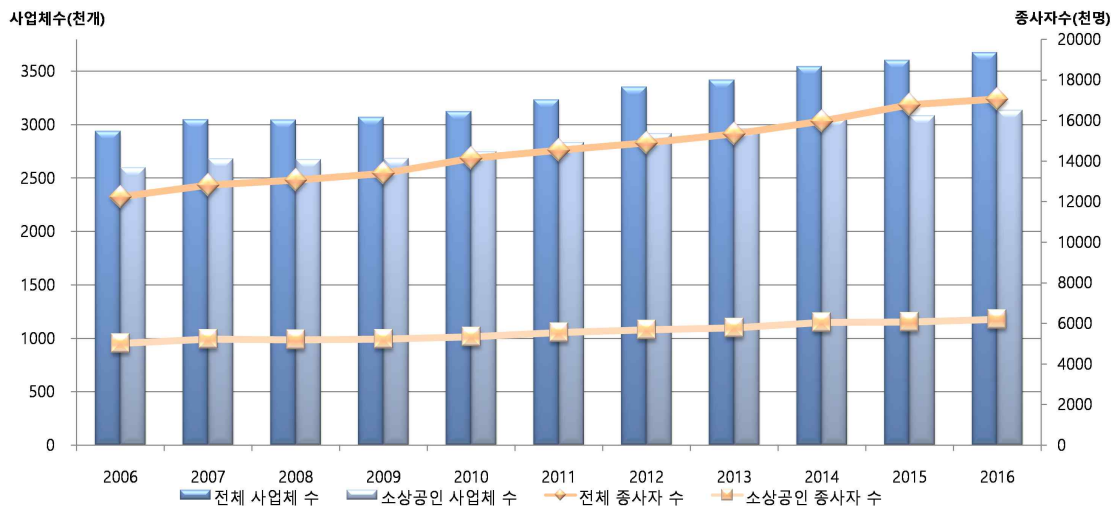
기업규모별 종사자수를 살펴보면 6,008,534명으로 중소기업 다음으로 많은 일자리를 창출하고 있음을 알 수 있다. 소상공인은 중소기업에 포함되지만 소상공인의 사업체수와 종사자수를 파악하기 위해 분리하여 나타냈다.



[그림 1] 기업규모별 종사자 수²⁾

2) 중소기업 중앙회, 2018년 중소기업현황(2016년 조사자료) 참고

아래 자료를 보면 2008년 이후 사업체 수와 종사자 수는 모두 꾸준히 증가하고 있으나 전체 사업체 수에서 소상공인 사업체 수가 차지하는 비중은 2008년 87.8%에서 2016년 85.3%로 감소하였고 전체 종사자 수에서 소상공인 종사자 수가 차지하는 비중은 2008년 39.7%에서 2016년 36.4%로 소폭 감소하였다.



[그림 2] 연도별 소상공인 현황³⁾

각 국가마다 소상공인의 범위를 다르게 정의하므로 국가 간 비교가 쉽지 않으나 비교가 가능한 지표로 자영업자 비율(Self-employment rate)이 있다. 자영업자는 소상공인과 의미가 비슷해 보이지만 다른 개념이다. 자영업자는 근로자를 고용하면서 사업을 영위하거나 자기 혼자 사업을 영위하는 사람이며 법인기업은 포함되지 않는다. 그리고 사업자 등록증 보유여부, 매출액, 상시근로자 규모와 무관하다. 우리나라의 자영업자 수는 OECD 회원국들 중 일부국가들의 자료를 정리한 아래 표와 같이 꾸준히 감소 추세이나 2018년 최근 조사결과를 보면 OECD 회원국들 중 4번째로 높은 비중을 차지하고 있음을 알 수 있다.

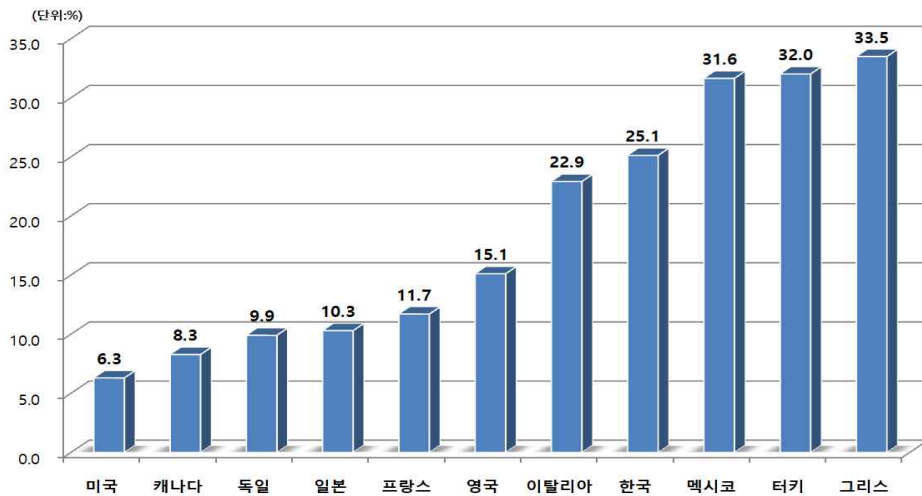
3) e-나라지표(<http://www.index.go.kr>) 참고

[표 4] OECD회원국의 자영업자⁴⁾ 비율 (2011년~2017년)

(단위 : %)

구분	2011년	2012년	2013년	2014년	2015년	2016년	2017년
그리스	36.2	36.6	37.0	36.0	35.0	34.1	34.1
터키	38.3	37.1	35.9	33.9	32.9	32.4	32.7
멕시코	33.7	33.0	32.9	32.2	32.1	31.7	31.5
한국	28.3	28.2	27.4	26.8	25.9	25.5	25.4
이탈리아	25.0	24.9	24.8	24.7	24.4	23.9	23.2
영국	14.2	14.6	14.6	15.2	15.0	15.4	15.4
프랑스	11.6	11.4	11.3	11.6	11.6	11.8	11.6
일본	11.3	11.8	11.5	11.4	11.0	10.5	10.4
독일	11.7	11.6	11.2	11.0	10.8	10.4	10.2
캐나다	9.0	8.9	8.8	8.8	8.6	8.6	8.3
미국	6.8	6.8	6.6	6.5	6.5	6.4	6.3

자료 : OECD (<https://data.oecd.org>)



[그림 3] OECD회원국의 자영업자 비율(2018년)⁵⁾

4) Self-employment : 자영업자 + 무급가족종사자

5) OECD (<https://data.oecd.org>)

2.3 소상공인 자금지원의 필요성

2017년 중소벤처기업부에서 264,337개 사업체를 대상으로 실시한 1인 창조기업 실태조사 자료를 보면 ‘1인 창조기업 경영 애로사항 1순위’ 조사항목에서 운영자금부족이 두 번째로 높았다.

[표 5] 1인 창조기업 경영 애로사항 1순위

(단위 : %)

경기불황	운영자금 부족	판로확보 (거래처)	기술력 부족	홍보 및 마케팅	해외시장 진출 및 개척
43.2	19.0	15.8	0.4	6.1	0.1
인력부족	인적 네트워크 부족	부가세 등 세금	업체 간 경쟁심화	시장 정보력 부족	기타
2.1	0.5	1.3	7.0	1.1	3.4

자료 : 통계청(2017)

신용보증재단중앙회에서 발표한 ‘2018 소상공인 금융실태조사 보고서’에서도 자금조달의 어려움을 알 수 있다. 각 지역신용보증재단을 방문하는 소상공인을 대상으로 설문조사한 자료에 의하면 창업과정에서의 애로사항과 경영과정의 애로사항으로 자금조달이 어려웠다는 응답이 가장 많았다. 이처럼 소규모 사업체는 규모가 큰 기업처럼 직접금융시장에서 채권이나 주식 발행 등으로 자금을 조달할 수 없으므로 자금 확보 경로가 다양하지 못하여 자금조달에 어려움을 겪고 있음을 알 수 있다.

[표 6] 소상공인 창업 애로사항

(단위 : 개, %)

구분	사례수	업종선택	입지선정	자금조달	행정절차	인력확보	지식부족	기타	
전체	2,589	5.3	17.0	61.7	4.2	6.1	3.7	2.1	
업종별	도소매업	885	4.9	14.5	66.9	3.3	4.1	4.2	2.3
	서비스업	548	3.5	19.9	58.4	6.2	7.5	2.0	2.6
	제조업	183	5.5	8.2	68.9	3.3	11.5	2.2	0.6
	음식숙박업	689	7.0	21.9	55.7	4.1	5.7	4.2	1.5
	기타업	284	6.0	12.7	61.3	4.2	7.0	5.3	3.5
성장단계별	개점초기	526	6.1	20.5	59.5	3.6	4.6	4.4	1.3
	성장기	784	4.6	17.4	63.1	4.2	5.4	3.1	2.3
	성숙기	896	5.3	17.3	59.8	4.8	8.2	2.9	1.8
	쇠퇴기	297	5.7	10.8	68.0	2.7	3.4	6.7	2.7
	폐업기	26	11.5	7.7	46.2	7.7	11.5	3.9	11.5
업력별	3년이하	1,071	5.5	20.0	59.9	4.3	5.2	3.7	1.4
	3년초과	1,515	5.2	14.8	63.0	4.2	6.7	3.7	2.6

자료 : 2018 소상공인 금융실태조사 보고서(신용보증재단중앙회)

[표 7] 소상공인 경영 애로사항

(단위 : 개, %)

구분	사례수	매출감소	자금조달	인력확보	업종쇠퇴	경쟁심화	정보/기술력부족	임대료상승	기타	
전체	3,725	17.1	33.0	15.7	3.7	21.6	1.6	6.2	1.1	
업종별	도소매업	1,293	16.1	38.8	12.4	4.3	21.5	1.5	4.6	0.8
	서비스업	768	14.7	29.0	16.2	4.3	25.1	2.1	7.4	1.2
	제조업	285	13.3	37.9	22.1	2.5	15.1	2.5	6.0	0.7
	음식숙박업	980	20.1	27.7	19.5	2.9	19.1	1.2	8.7	0.9
	기타업	399	20.3	31.6	11.8	3.3	25.8	1.8	3.0	2.5
성장단계별	개점초기	724	14.0	38.5	13.5	1.8	22.9	2.5	5.9	0.8
	성장기	1,131	11.3	37.1	17.8	2.0	23.4	1.8	5.7	1.0
	성숙기	1,275	15.8	29.7	19.1	4.0	22.4	0.9	7.1	1.1
	쇠퇴기	452	37.4	25.2	5.1	9.1	15.3	2.0	4.9	1.1
	폐업기	39	20.5	23.1	10.3	12.8	15.4	0.0	12.8	5.1
업력별	3년이하	1,489	12.6	37.1	16.1	2.2	22.6	2.0	6.5	0.9
	3년초과	2,233	20.1	30.3	15.4	4.7	21.0	1.4	6.1	1.2

자료 : 2018 소상공인 금융실태조사 보고서(신용보증재단중앙회)

제 3 장 선행연구

3.1 기업부실의 개념⁶⁾

기업부실(corporate financial distress)은 부실의 상황이 다양하고 복잡하므로 한마디로 정의하기 어려우나 일반적으로 경제적 실패, 지급불능, 파산 등을 포괄하는 개념으로 사용된다.

경제적 실패(economic failure)란 총수익이 총비용보다 낮은 경우, 투자 수익률이 자본조달비용보다 낮은 경우 등 주로 회사의 수익성저하가 원인이 되어서 나타나는 경제적 문제를 의미한다. 그러나 경제적 부실만으로는 그 기업이 곧 무너질 기업이라고 볼 수는 없다. 그 이유는 경제적 부실현상이 존재한다 하더라도 당장에 갚아야 할 부채가 없고 부채를 상환할 수 있는 능력이 유지된다면 계속 존속할 수 있기 때문이다.

지급불능(insolvency)은 유동성이 부족하여 만기가 도래한 채무를 상환하지 못하는 기술적 지급불능 (technical insolvency)과 총자산가치가 총 부채가치보다 적어서 실질순자산가치가 마이너스로 되는 실질적 지급불능(real insolvency)으로 구분된다. 기술적 지급불능은 파산의 직접적인 원인이 될 수 있지만 일시적 현상인 경우가 많다. 기술적 지급불능이 있더라도 만기가 도래한 채무를 갚는 데 필요한 유동성이 즉시 뒷받침 될 수 있다면 기업이 정상적인 상황으로 회복될 수 있기 때문이다. 그러나 실질적 지급불능은 기술적 지급불능과는 다르게 그 원인이 일시적인 유동성의 부족이 아니라 만성적인 결손누적에 있는 것이 일반적이다.

6) 박정식, 신동령, 2016. 경영분석. 다산출판사.제10장 제1절 참고

정상기업이 부실기업이 되는 과정은 수익성 감소, 지급능력 저하, 법률적 파산의 3단계를 거쳐 진행된다고 볼 수 있다. 그 과정에서 기업부실을 미리 예측하여 피해를 최소화하기 위해서는 부실 전 어떤 징후들이 나타나는지 파악하고 있어야 한다. 아래 표와 같이 재무제표의 각 항목들을 분석하여 부실을 예측하는 방법은 자료의 수집이 비교적 쉬우므로 널리 이용되는 방법이다. 하지만 부실의 징후가 재무제표에 즉각 반영되기 어려우며 재무제표에 반영되지 않는 비재무적 정보들이 존재하므로 비재무적징후들을 파악하여 분석할 필요가 있다. 비재무적 징후는 기업의 환경에 따라 다양하나 부실기업에서 발견할 수 있는 주요변화로는 생산(가동률 하락, 근무시간 단축, 전기소모량 감소, 제품 불량률 급증 등), 구매(매입처의 이유 없는 변경, 긴급한 구입의 빈발 등), 판매(경쟁상품 출현, 재고누적, 출혈판매, 대금결제지연 등), 경영자(약속위반, 행선지불명 등), 직원(유능한 인력 퇴사, 근무기강 해이 등)의 변화가 있다.

[표 8] 재무적 부실징후

손익계산서	대차대조표	현금흐름표
-매출액의 감소 -매출원가, 판매비, 관리비 등 비용 급증 -매출총이익 적자 -영업이익 적자 -과다한 금융비용 부담 -이익률의 현저한 하락	-현금과 현금성자산의 부족 -매출채권 및 재고자산 급증 -단기차입금 과다 -부채 과다 -불리한 조건의 신규차입 급증 -타인자본의존도 심화 -이월결손금 과다 -자본잠식 -결손의 누적	-현금 및 현금성자산 감소 -영업 활동, 현금 흐름의 부족 또는 적자 지속 -단기부채조달의 증가 -장기부채조달의 감소 -과다한 배당금의 지급

3.2 부실예측 선행연구

부실예측에 대한 체계적인 연구는 Beaver(1966)에 의해 시작되었다. Beaver는 1954년부터 1964년까지의 기업을 조사하여 회사채 채무 불이행, 부도수표발행, 우선주 배당금 미지급, 파산 등에 해당하는 79개의 부실기업과 업종 및 자산규모가 유사한 79개의 정상기업을 대응시켜 표본으로 선정하였다. 그리고 프로파일분석과 이원분류검정을 실시하여 재무비율의 기업부실 예측능력을 평가하였다.⁷⁾ 단일변량분석모형은 기업의 파산을 예측하기 위해 하나의 재무비율을 사용하므로 부실의 예측변수를 쉽게 찾을 수 있다. 그러나 부실로 예측된 재무비율 간의 상관관계를 알 수 없고 어느 정도의 예측능력을 가지고 있는지 알기 어렵다. 또한 선정된 재무비율에 따라 예측결과가 달라질 수 있다는 문제점이 있다.

단일변수만을 고려하는 단일변량분석의 한계점을 보완할 수 있는 다변량판별분석(Multiple Discriminant Analysis:MDA)은 여러 변수들 간의 관계를 동시에 고려하여 2개 이상 집단의 차이를 분석하는 방법이다. Altman(1968)은 기업의 부실예측에 처음으로 다변량판별분석기법을 활용하였고 부실예측을 위한 여러 개의 재무비율 중 가장 영향력이 높은 재무비율 5개를 선택하여 부실을 예측할 수 있는 Z-Score모형을 개발하였다. 그는 1946년에서 1956년 사이에 파산신청을 한 33개 기업을 부실기업의 표본으로 선정하였고 이 기업들과 업종이 동일하고 자산규모가 유사한 정상기업 33개를 정상기업의 표본으로 선정하였다. 그리고 연구목적에 적합한 것으로 판단되는 22개의 재무비율을 선택하고 그 중 수익성, 유동성, 레버리지, 지급능력, 활동성을 대표하는 5가지 재무비율(이익잉여금/총자산, 운전자본/총자산, 영업이익/총자산, 자기자본의 시장가치/

7) Beaver, W. H., 1966. Financial Ratios as Predictors of Failure, Journal of Accounting Research, pp.71-111.

총부채, 매출액/총자산)을 판별변수로 선정하여 다변량판별분석을 실시하였다. 그러나 판별식에 적용할 변수의 객관적인 선정기준이 없으며 예측 오차가 크게 발생할 가능성이 있다는 한계점이 있다.8)

다변량판별분석의 한계점을 극복하기 위해 Ohlson(1980)은 로짓분석(logit analysis)을 부실기업 예측에 활용하였다. 그는 1970년부터 1976년의 기간 중 파산을 신청한 상장 제조업 105개를 부실기업의 표본으로, 2058개의 비도산기업을 정상기업의 표본으로 선정하여 로짓분석을 실시하였다. 그 결과, 기업의 규모, 운전자본/총자산, 총부채/총자산, 영업활동현금흐름/총부채 등이 유의한 영향을 주는 것을 알 수 있었다. 로짓모형은 설명변수가 정규분포하며 집단 간 분산·공분산행렬이 동일해야 한다는 통계적 조건을 요구하지 않으며 설명변수들의 영향력을 서로 비교할 수 있고 도산확률을 구할 수 있으므로 해석이 쉽다는 이점이 있다.

[표 9] 다변량판별분석, 로짓모형을 이용한 선행연구

연구자	연구내용	연구표본	지표	분석기법
강치형 신해수 (2015)	회원제 골프장 기업의 부도 예측 모형개발	2009년~2014년 부실기업:22개 정상기업:22개	고정장기적합률, 부채비율, 자기자본비율, 차입금의존도, 입회보증금의존도, 순운전자본비율, 유동비율, 당좌비율, 자기자본순이익률, 총자산순이익률, 매출원가율, 매출액영업이익률, 매출액경상이익률, 금융비용대비매출액, 이자보상비율, 자기자본회전율, 유형자산회전율, 자기자본증가율, 총자산증가율, 매출액증가율, 총자본투자효율, 유형자산투자효율, 부가가치율	다변량판별분석, 로지스틱 회귀분석
김상봉 조경준 (2011)	부도예측모형을 이용한 기업부실화의 원인분석	2006년~2008년 부실기업:1,617개 정상기업:1,107개	자기자본비율, 유동비율, 당좌비율, 차입금의존도, 매출액규모, 총자산규모, 부채상환계수, 총자본순이익률, 총자본경상이익률, 매출액경상이익률, 금융비용대비매출액, 매출원가율, 이자보상비율, 총자본회전율, 재고자산회전율, 총자본증가율, 유형자산증가율, 매출액증가율, 순이익증가율, 총자본투자효율, 부가가치율	로짓모형

8) Altman E. I., 1968. Financial Ratios, Discriminant Analysis and the Prediction Corporate Bankruptcy, Journal of Finance Vol.23, pp.589-609.

윤일현 (2011)	금융회사의 도산예측모형에 관한 연구	1998년 부실기업:16 정상기업:314개	자기자본/총여신, 총여신/총자산, 영업비용/영업수익, ROA, ROE, 현금및은행예치금/차입금및고객예수금, 총자산의자연로그값, 외화차입금/총외화차입금및예수금	로짓분석
전현우 정용화 신동휴 (2011)	상장폐지기업의 부실예측모형에 관한 연구 -거래소시장을 중심으로-	2007년-2010년 부실기업:31개 정상기업:31개	성장성비율(총자산증가율, 유형자산증가율, 유동자산증가율, 매출액증가율), 수익성비율(총자본영업이익율, 매출원가율 등 20가지), 자본구조비율(유동비율, 유보액/납입자본비율, 총CF대부채비율, 총CF대총자본비율, 총CF대매출액비율, 현금비율, 매출채권/매입채무비율, 순운전자본/총자본비율, 유보액/총자산비율, 당좌비율), 회전율(총자본회전율, 순영업자본회전율, 자본회전율, 순운전자본회전율, 경영자본회전율, 비유동자산회전율, 유형자산회전율, 매출채권회전율, 매입채무회전율)	다변량판별분석
유승규 박정로 최재규 김재준 (2009)	다변량판별분석(MDA)기법을 이용한 외부감사 건실 기업의 부도 예측 모형개발에 관한 연구	2008년 부실기업:33개 정상기업:33개	재고자산증가율, 자기자본증가율, 부채비율, 유동비율, 차입금의존도, 자기자본구성비율, 매출액순이익률, 매출채권회전율	다변량판별분석
권영린 (2002)	호텔기업의 부실예측모형에 관한 연구	1998년-2000년 부실기업:9개 정상기업:9개	자기자본비율, 이자보상비율, 유동비율, 총자본경상이익율, 매출액경상이익율, 총자산회전율, 고정자산회전율, 객실점유율, RevPar, 평균객실요금, 각지표들의 변동성	로짓분석
김용덕 이인자 (2002)	국내정보통신산 업의 부실분석 및 예측모형에 관한 연구	1998년-2001년 부실기업:22개 정상기업:22개	총자산성장률, 순운전자본구성비율, 총자산영업이익률, 매출액성장률, 고정자산회전율, 고정장기적합률, 당좌비율, 부채비율, 소재지, 영업연도	다변량판별분석, 로짓분석
강종만 홍성희 (1999)	부실예측모형의 적합성 분석	1991년-1997년 부실기업:65개 정상기업:65개	종업원1인당부가가치증가율, 금융비용대총비용비율, 1주당매출액, 이자보상배율, 유보액대납입자본비율, 매입채무회전율, 매출채권회전기간, 종업원1인당인건비, 총위험, 주가초과수익률	다변량판별분석(Z점수모형, ZETA 모형), 로짓분석

최근에는 과학기술이 점차 발달하게 되면서 과거의 통계학적 기법에 비하여 예측정확도가 향상된 머신러닝(Machine Learning)의 활용도가 증가하였다. 머신러닝을 활용한 연구는 금융, 건설, 공학, 화학 등 다양한 산업 분야에서 활용되고 있는 추세이다. 머신러닝은 인공지능(Artificial Intelligence)의 하위분야로 컴퓨터가 학습할 수 있는 알고리즘과 기술을

개발하는 영역이며 컴퓨터는 설계된 알고리즘을 따라 빅데이터에 담긴 정보를 학습하여 예측력 제고에 유용한 변수를 귀납적으로 학습하고 결론을 도출한다. 머신러닝은 변수간의 관계와 같은 사전적 지식을 요구하지 않으며 주어진 알고리즘 틀을 이용하여 데이터를 학습하고 규칙을 찾을 수 있다. 이러한 이점들로 인하여 머신러닝을 활용하여 대출신청자의 신용도를 판단하는 사례가 증가하고 있다. 일반 은행들이 소득, 금융거래 내역, 연체여부 등 제한적인 변수에 의한 신용평가 모델을 적용하는 반면, 최근 핀테크 기업들은 빅데이터를 활용하여 수백 개의 변수를 검토하고 최적의 평가 모델을 적용한다.

[표 10] 머신러닝을 이용한 선행연구

연구자	연구내용	연구표본	지표	분석기법
이주희 동학림 (2018)	소상공인의 자금공급 확대를 위한 빅데이터 활용방안 연구	2009년~2018년, 부실기업:2,011개 정상기업:15,105개	업력, 신용등급, 규모, 매출액, 업종별부도율, 주거인구, 직장인구, 유동인구, 세대수, APT수, 경쟁점포, 업종매출	데이터마이닝 (관별분석)
표수진, 고형진, 손범호, 김호기, 이재욱 (2018)	머신러닝 기반 저축은행의 부실 징후 예측 및 업권 위험지표 지수 개발	2001년 이후 정상 및 부실 저축은행	고정이하여신비율, 기본자본비율, 총자산순이익률, 수지비율, 예대율, 차입금의존도, 총자산규모, 부채상환계수, 총자산순이익률, 영업수익대비영업비용, 총자산대비율, 유동성부채대비순대출, 유형자산비율	나이브베이즈, 의사결정나무, SVM, K최근접이웃, 로지스틱회귀분석, 다층퍼셉트론, 랜덤포레스트, 그라디언트부스팅, 익스트림그라디언트부스팅
김시중 (2016)	부실예측모형의 예측력 비교	2015년 서울과 제주에 위치한 특1급 호텔 정상기업:12개 부실기업:9개	부채비율, 차입금의존도, 유동비율, 영업이익대비이자보상비율, 매출액영업이익률, 매출액순이익률, 총자산순이익률, 자기자본순이익률, 영업현금흐름, 매출액증가율, 영업이익증가율, 당기순이익증가율, 총자산증가율, 총자산회전율	다변량관별분석, 로짓분석
이수현, 박정민, 이형용 (2015)	데이터마이닝 기법을 활용한 비외감기업의 부실화 유형 분석	2004년, 부실기업 99개	유동비율, 부채비율, 총차입금의존도, 이자보상배율, 총자본경상이익률, 매출액경상이익률, 총자산회전율, 총자산증가율, 매출액증가율, 현금흐름대유동부채비율	자기조직화지도 (SOM)
윤종식 권영식	SVM을 이용한 소상공인 부실	28개월간의 카드가맹점 정보	6개월동안카드거래중단개월수, 최대매출월과최소매출월사이기간, 기준시점까지	SVM, 인공신경망, 의사결정나무, 다변

(2007)	예측모형	정상기업:5000개 부실기업:5000개	의지속거래기간,매출발생개월매출평균,최고액과최소액의차이금액,6개월최소카드매출액,최소건수,3개월건수평균의분위수,3개월거래의변화방향,3개월거래의변화정도,중분류업종부도율	량관별분석,로지스틱회귀분석
민재형, 정철우 (2007)	기업도산예측을 위한 이진 분류 기법의 개발	2001년~2004년, 부실기업:1271개 정상기업:1271개	총자본투자효율,총자산증가율,유동부채대총자산,자기자본비율,총자산회전율,감가상각비,순금융비용,손익분기점율,인건비	유전알고리즘(GA), 의사결정나무,인공 신경망,판별분석, 로지스틱회귀분석
이영찬 (2004)	인공신경망과 Support Vector Machine의 기업부도예측 성과비교-Support Vector Machine의 유용성을 중심으로-	1999년~2002년, 국내 비외감 정보통신기업 정상기업:258개 부실기업:258개	금융비용대매출액,매출액순이익율,이자보상비율,총자본경상이익율,순운전자본비율,유동비율,자기자본비율,차입금의존도,매출채권회전율,현금흐름대총부채	인공신경망,SVM
김진백 이준섭 (2000)	인공지능기법을 이용한 중소기업 부도예측에 있어서 현금흐름변수의 유용성 검증	1995년~1997년, 외감대상 제조업 정상기업:103개 부실기업:103개	매출액,총자산증가율,총자본경상이익률, 고정장기적합율의 역, 유동부채대총자본, 자기자본비율,이자지급후CF/금융비용,외부자금조달후CF/금융비용	ANN,CBR,MDA

제 4 장 연구 방법

4.1 랜덤 포레스트(Random Forest)

머신러닝(Machine learning)기법 중 하나인 랜덤 포레스트(Random forest)는 의사결정나무(decision tree) 기반의 앙상블 모형이다. 앙상블이란 단일모형들보다 더 좋은 예측성능을 얻기 위해 다수의 알고리즘을 사용하는 기법이다. 랜덤 포레스트는 각 의사결정나무로부터 얻은 값을 평균 내어 회귀분석하거나 다수결 투표의 방법으로 분류(Classification)에 이용할 수 있다. 랜덤포레스트는 2001년 레오 브레이먼(Leo Breiman)의 연구에서 처음으로 도입되었으며, 임의노드최적화(Randomized node optimization)와 배깅(Bagging)으로 의사결정나무(decision tree)의 문제점을 보완한 기법이다. 의사결정나무(decision tree)는 의사결정 규칙을 나무구조로 나타내어 해석이 용이하다는 장점이 있으나 결과 값의 변동 폭이 크다는 단점을 가지고 있다. 훈련데이터를 적절하게 입력하지 못할 경우 편향(bias)과 분산(variance)이 커지고 복잡한 의사결정나무를 만들어 과적합(overfitting)문제가 생길 수 있고, 진행과정에 오류가 발생하면 다음 단계로 오류가 전파되는 문제가 있다.

랜덤포레스트는 배깅(bagging)과 임의노드최적화(Randomized node optimization)의 무작위성(randomness)에 의해 비슷한 의사결정나무들이 만들어지지 않게 되고 모형들 간의 상관성을 낮춤으로써 예측성능을 향상시킬 수 있다.

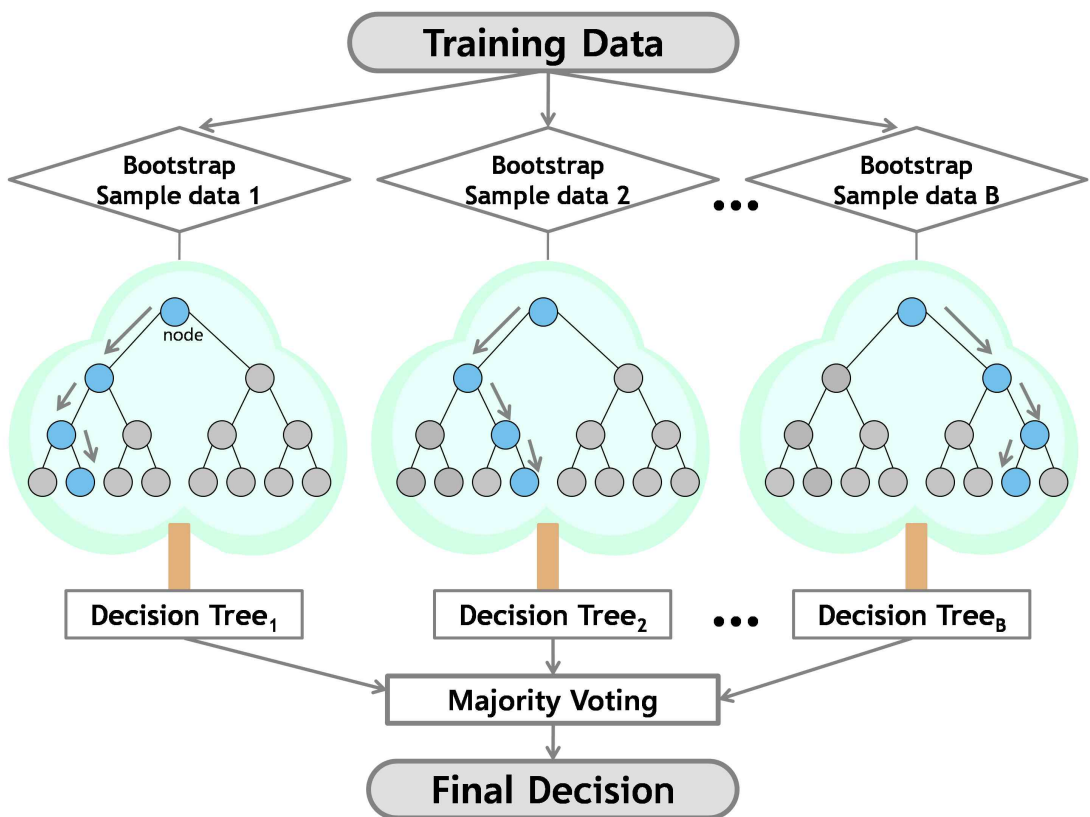
배깅(bagging)은 bootstrap aggregating의 줄임말이며 부트스트랩(bootstrap)방식으로 추출한 훈련데이터로부터 부트스트랩 표본데이터를 B 번 생성하여 각 데이터마다 예측모형을 만든 후 예측결과를 조합하여 하

나의 최종예측모형을 만드는 방법이다. 배깅(bagging)은 중복을 허용하여 동일한 크기의 표본을 복원 추출하는 부트스트랩방식에 의해 예측 오차의 분산(variance)을 줄여 예측력을 향상시키고 불안정한 분류모형들을 안정적으로 만들 수 있다.

임의노드최적화(Randomized node optimization)는 노드 분할함수의 매개변수 최적값을 구하는 방법이다. 데이터들은 의사결정나무의 각 노드(node)에서 분할 함수(split function)에 의해 분할되며 각 노드마다 무작위로 다른 변수가 선택되어 분할된다. 각 분할에 p 개의 예측변수를 모두 고려하면 만들어진 나무들이 유사하여 단일 모델에 비해 분산을 크게 줄이지 못하지만 각 분할에서 무작위로 선택된 p 보다 작은 m 개의 예측변수를 사용하여 나무들의 상관성(correlation)을 줄일 수 있다. 유사한 나무가 생성되지 않게 하기 위한 m 값으로 분류에서는 $m = \sqrt{p}$, 회귀에서는 $m = \frac{p}{3}$ 를 일반적으로 사용한다. 단일 의사결정나무는 노이즈에 대해서 민감하지만, 나무들이 서로 상관화되지 않는다면 여러 나무들의 평균값은 노이즈에 대해 더욱 견고해지며 상관관계의 감소는 랜덤포레스트의 일반화 성능을 향상시킨다.

랜덤포레스트의 성능평가는 오분류율로 파악할 수 있다. 부트스트랩에서 데이터를 추출할 때 추출되지 않은 데이터를 OOB(Out-Of-Bag)데이터라 한다. OOB(Out-Of-Bag) 데이터를 알고리즘 수행 후 얻어진 나무 모형들에 적용시켜 오분류율을 구할 수 있다. 오분류율은 예측에 실패한 데이터의 비율이며 오분류율이 낮을수록 예측이 정확하게 된 것이다. 랜덤포레스트는 의사결정나무모형보다 해석이 어려우나 변수 중요도를 측정하여 독립변수들 간의 영향력을 비교할 수 있다. 변수 중요도는 변수가

정확도의 개선과 노드 불순도의 감소에 얼마나 기여하는지를 계산하여 파악할 수 있다. 그리고 변수의 중요도를 검토하여 예측 과정에 적절하지 않은 어떤 기능을 삭제할지를 선택할 수 있다. 랜덤 포레스트는 매우 큰 데이터집합에도 잘 작동하며 복잡한 통계적 가정이 없기 때문에 데이터의 전처리도 단순하고 다양한 종류의 변수를 다룰 수 있다는 이점이 있다. 랜덤 포레스트의 구조를 도식화하면 아래 그림 4 와 같다.



[그림 4] 랜덤포레스트(Random Forest)의 구조

랜덤 포레스트 알고리즘 (Random forest algorithm)⁹⁾

- (1) 훈련데이터 $L = \{(x_i, y_i), i = 1, \dots, n\}$ 을 정의한다. 여기서 x_i 는 입력변수 벡터이고 y_i 는 목표변수이다. 입력변수를 p 개라 가정하면, $x_i = (x_{i1}, \dots, x_{ip})'$ 이고 입력변수를 랜덤 추출한 벡터 x_i^* 는 x_i 에서 m 개의 변수로 구성된 입력변수집합이다. ($m \ll p$)
- (2) L 로부터 B 개의 부트스트랩 데이터 L_1, \dots, L_B 를 만든다.
- (3) L_b 를 이용하여 $C_b(x)$ 를 생성한다. 단, $C_b(x)$ 를 생성할 때 매 중간노드마다 x_i 를 사용하지 않고 x_i^* 를 사용하여 의사결정나무를 생성한다.
- (4) $b = 1, \dots, B$ 에 대하여 단계 (3)의 과정을 반복 수행한다.
- (5) B 개의 분류기를 결합시켜 최종예측모형 $\hat{f}(x)$ 를 만든다. 여기서 x 는 예측하고자 하는 관측치의 입력변수 벡터 값이다.

$$\hat{f}(x) = \operatorname{argmax}_j \left\{ \sum_{b=1}^B I[C_b(x) = j], j = 1, \dots, J \right\}$$

4.2 로지스틱 회귀분석(Logistic Regression)¹⁰⁾

로지스틱 회귀분석(Logistic Regression)은 통계학자인 D.R.Cox가 1958년에 통계적 방법으로 개발한 모델로서 하나 이상의 독립변수가 있는 데이터집합을 분석하기 위한 방법으로 의학, 사회과학 등 많은 분야에서 사용되어온 통계기법이다. 판별분석과 달리 독립변수의 연속성이나 정규분포성에 대한 전제가 필요하지 않으며 종속변수가 수치화하기 힘든 범주형 데이터를 다룰 때 유용하게 쓰인다.

로지스틱 회귀분석은 독립변수의 특성에 따라 컨디셔널 로지스틱 회귀

9) 장영재, 김현중, 조형준, 2016. 데이터마이닝. 한국방송통신대학교출판문화원. 제4장 참고

10) 이훈영, 2012. 연구조사방법론, 청람. 제3부 제20장 참고

분석과 일반적인 로지스틱회귀분석으로 구분된다. 일반적인 로지스틱 회귀분석은 종속변수의 범주 수에 따라 크게 두 가지 종류가 있으며, 종속 변수가 2개의 범주를 가지는 이항 로지스틱 회귀분석(binomial logistic regression)과 3개 이상의 범주를 가지는 다항 로지스틱 회귀분석(multinomial logistic regression)으로 구분된다. 이항 로지스틱 회귀분석은 기업의 도산예측, 대출금의 연체여부 등 범주가 2개일 경우 사용되며, 본 연구에서도 이항 로지스틱 회귀분석을 이용하여 부실여부를 예측하고자 한다.

이항 로지스틱 회귀분석은 종속변수가 두 가지 값만 취하는 이항 범주형 변수일 경우 종속 변수와 독립 변수간의 관계를 함수로 나타내어 종속변수 값을 예측하는데 사용되는 대표적인 통계분석 방법이다. 종속 변수를 독립 변수의 선형 결합으로 설명하는 점에서는 선형 회귀 분석과 유사하지만 종속 변수가 범주형 데이터를 대상으로 한다는 점에서 선형 회귀 분석과는 다르다. 일반적인 회귀분석에서 대부분의 경우 종속변수는 음의 무한대(-∞)에서 양의 무한대(+∞)에 해당하는 값을 가지는 것을 전제로 한다. 그러므로 종속변수 값으로 0과 1처럼 범주형 값을 가지는 자료를 일반적인 회귀분석으로 분석하는 것은 적합하지 않으므로 종속변수의 값을 변환하는 과정이 필요하다. 분석을 위해 범주형 종속변수를 확률 값으로 변환하고 종속변수(y)가 특정한 값을 가질 확률(p)과 반대 확률(1-p)의 비인 오즈(odds)값 $\left(\frac{p}{1-p}\right)$ 을 구한다. 그리고 확률의 오즈(odds)값 $\left(\frac{p}{1-p}\right)$ 에 자연상수 e를 밑으로하는 자연로그를 취한 로그오즈값 $\log_e\left(\frac{p}{1-p}\right)$ 을 종속변수로 하여 분석한다. 이처럼 로지스틱 회귀분석은 종속변수를 변환한 회귀분석의 일종이며 로그오즈를 이용한 회귀모형을 로지스틱 회귀모형(logistic regression model)이라 하고 식 (1)은 아래와 같

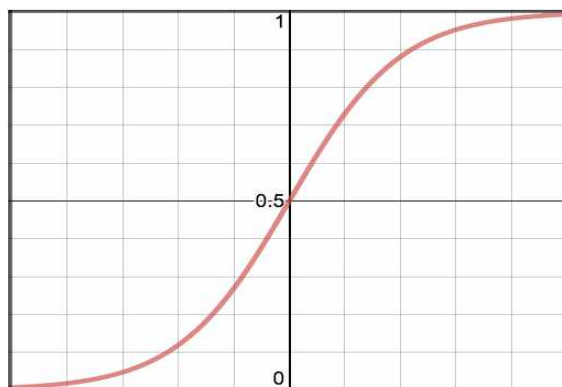
이 나타낼 수 있다.

$$\log_e \left(\frac{p_i(y_i = 1)}{1 - p_i(y_i = 1)} \right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_m x_{im} \quad (1)$$

이러한 로지스틱 회귀모형을 종속변수(Y)가 특정 범주 값인 1을 가지는 확률 p에 대하여 정리하면 아래의 식 (2)와 같은 로지스틱 회귀방정식을 구할 수 있다.

$$p = P(Y = 1 | X = x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \quad (2)$$

위 식을 그래프로 표현하면 아래와 같은 곡선의 형태이며 이를 로지스틱 회귀곡선이라 한다.



로지스틱 회귀모형의 계수는 우도 추정법(MLM:Method of Maximum Likelihood)을 이용하여 확률값인 우도 함수(likelihood function)를 최대로 하는 회귀계수를 구할 수 있다. 그리고 로지스틱 회귀모형의 적합도는 우도비 검정 (likelihood ratio test), 유사결정계수(Pseudo - R^2), 적중률 (hit-ratio)로 검정할 수 있다.

4.3 평가지표

본 연구에서는 분류모델의 성능을 평가하기 위해 결과 값을 혼동행렬 (Confusion Matrix)로 정리하고 평가지표로는 정확도(Accuracy)와 재현율(Recall)을 사용하였다.

[표 11] Confusion Matrix¹¹⁾

		Actual	
		Positive	Negative
Predicted	Positive	True Positive(TP)	False Positive(FP)
	Negative	False Negative(FN)	True Negative(TN)

각 모델의 예측 결과는 혼동행렬(Confusion Matrix)구조로 표현될 수 있다. 이 행렬에는 네 가지 범주가 있는데 행은 예측 값을 나타내며 열은 실제 값을 나타낸다. True Positive는 실제 값과 예측 값이 모두 positive로 예측한 것을 의미하고 False Negative는 실제 값은 Positive이지만 Negative로 예측한 것을 의미한다. 그리고 False Positive는 실제 값은 Negative지만 예측 값은 Positive인 것을 의미하고 True Negative는 실제 값과 예측 값이 모두 Negative인 것을 의미한다.

정확도(Accuracy)는 전체 데이터에서 Positive를 Positive로 예측하고 Negative를 Negative로 맞게 예측한 결과를 더한 비율이다. 본 연구에서는 전체 예측결과 중에서 부실기업을 부실기업으로 맞게 예측한 것과 정상기업을 정상기업으로 맞게 예측한 것을 더한 비율을 나타낸다.

11) Davis, J. & Goadrich, M., 2006. The relationship between Precision-Recall and ROC curves, In Proceedings of the 23rd international conference on Machine learning, pp.233-240.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (3)$$

재현율(Recall)은 민감도(sensitivity)라고도 하며 실제 Positive중에서 Positive로 맞게 예측한 것의 비율이다. 본 연구에서는 실제 부실기업 중에서 부실기업으로 맞게 예측한 비율을 나타낸다.

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

제 5 장 소상공인 정책자금 부실예측

5.1 표본기업의 선정

분석대상이 되는 표본기업은 2016년부터 2017년 4월까지 제조업에 종사하는 소상공인을 대상으로 하는 소상공인 정책자금을 지원받은 조선기자재 분야 제조업을 영위하는 소상공인이다. 자금의 상환은 대출실행 후 다음 달부터 원리금을 상환하는 것이 아니라 2년간 소액의 이자만 납부하는 거치기간이 있다. 거치기간동안에는 원금상환의 부담이 없어서 사고 발생율이 낮으므로 동일한 조건에서 부실을 파악하기 위해 대출실행 일로부터 2년이 지나 원금 상환이 시작된 기업만을 대상으로 하였다. 그리고 연구를 위해 소상공인 정책자금을 지원받은 기업 중 사고가 발생하지 않은 180개 기업을 정상기업으로, 사고기업으로 분류된 65개 기업을 부실기업으로 선정하였다. 사고기업은 기한의 이익 상실사유가 있는 기업이며 구체적인 내용은 다음과 같다.

- 1) 대출금이자 등을 상환하여야 할 때부터 지체일수가 60일에 이른 때
- 2) 분할상환금 또는 분할상환원리금을 상환하여야 할 때부터 지체횟수가 3회 이상 이른 때
- 3) 「채무자 회생 및 파산에 관한 법률」에 의한 회생절차·파산절차·개인회생 절차 개시의 신청이 있거나 청산에 들어간 때
- 4) 조세공과에 관하여 납기 전 납부고지를 받거나 어음교환소의 거래정지처분을 받을 사유가 발생한 때
- 5) 채무자가 제공한 담보물건 및 자가사업장(임차사업장의 경우에는 임차보증금)에 대하여 압류, 가압류, 가처분 결정 또는 경매신청이 있거나 체납처분 착수가 있는 때
- 6) 한국신용정보원의 “신용정보관리규약”에 따른 연체·대위변제·대지급·부도정보(이상 관련된 정보 포함), 금융질서문란정보 및 체납정보 등의 공공정보 등록사유가 발생한 때

- 7) 신용회복위원회에 신용회복지원을 신청한 때
- 8) 폐업, 조업중단, 대표자 도피 등의 사유로 기업의 지속적인 영업이 곤란한 때
- 9) 법인가업의 대표자(실제경영자 포함), 개인기업의 공동경영자(실제경영자 포함)에 대하여 위 제3호 내지 제7호의 사유가 발생하여 이로 인하여 본인의 신용상태가 현저히 악화된 때
- 10) 제3호 내지 제8호에 준하는 사유로 신용상태가 크게 악화된 것이 객관적으로 확인된 때
- 11) 채무자에 관하여 그 외 기타사유로 채권보전에 현저한 위험이 예상될 때

5.2 변수 선정

본 연구에 이용되는 변수는 소상공인 정책자금지원기업의 재무적 정보와 비재무적 정보 중 기업부실과 연관성을 보이는 아래 지표들을 변수로 선정하였다.

[표 12] 분석에 사용된 변수

재무적 변수	설명
Sales	매출액
SalesbyWK	1인당 매출액(매출액/상시근로자 수+1)
Loan2Debt	부채대비 대출신청금액
Debt	총부채
NetProfit	당기순이익
OpProfit	영업이익
Asset	총자산
ROE	ROE(당기순이익/자본총계)
Interest	이자보상배율(영업이익/이자비용)
Debratio	부채비율(부채총계/자본총계)
Receivable	매출채권회전율(매출액/평균매출채권)
Inventory	재고자산회전율(매출액/평균재고자산)
AssetTurnover	총자산회전율(매출액/총자산)

비재무적 변수	설명
LoanF2	2금융권 대출 여부
Worker	상시근로자 수
CB	CB(Credit Bureau)점수 : 신용조회회사에서 제공하는 개인신용점수 (개인대출정보, 개인보증실적, 카드실적정보, 연체정보, 신용거래개설정보, 신용정보기관 채무불이행정보, 한국신용정보원 현금서비스 정보 등의 종합점수)
CBrate	CB등급 : 개인신용점수를 등급화한 것
SOHO	SOHO점수(CB점수+기업점수) : 신용조회회사에서 제공하는 기업신용점수 (사업자개요, 신용공여정보(은행, 비은행 대출채권정보), 연체 및 신용도 판단정보(신용불량정보), 카드실적정보, 한국신용정보원 현금서비스 정보 등의 종합점수)
SOHORate	SOHO등급 : SOHO점수를 등급화한 것
Busi	사업성평가점수(심사자가 사업성, 시장성, 경영능력을 평가한 점수)
Busirate	사업성등급 : 사업성평가점수를 등급화 한 것
Firmrate	기업평가등급 : SOHO등급 + 사업성평가등급
Life	업력 : 개인기업은 사업자등록증명상의 개업일로부터, 법인기업은 법인등기사항전부증명서상의 법인설립등기일로부터 신청접수일까지의 기간

5.3 실증분석

랜덤포레스트(Random Forest)와 로지스틱회귀(Logistic Regression)모형을 이용하여 자료를 분석하였으며 변수의 구성에 따라 각각 3가지 모형을 만들어 실험을 실시하였다. Model 1은 재무적 변수만 반영한 것이고 Model 2는 비재무적변수만 반영한 것이다. 그리고 Model3은 Model 1과 Model 2에서 가장 기여도가 높은 상위 5개의 재무적변수와 비재무적변수 5개씩을 선정하여 반영한 것이다. 재무적 변수 5가지는 기여도순으로 영업이익(OpProfit), 당기순이익(NetProfit), 매출액(Sales), 총자산회전율(AssetTurnover), 부채비율(DebtRatio)이며 비재무적 변수 5가지는 소호점수(SOHO), CB점수(CB), 업력(Life), 사업성점수(Busi), 상시근로자 수(Worker)이다.

5.3.1 Random Forest 모형의 평가

[표 13] Random Forest 모형의 예측 결과

Random Forest			Actual		Accuracy	Recall (Sensitivity)
			정상(0)	부실(1)		
Model 1	Predicted	정상(0)	34	12	63.2%	14.3%
		부실(1)	9	2		
Model 2	Predicted	정상(0)	42	13	75.4%	7.1%
		부실(1)	1	1		
Model 3	Predicted	정상(0)	39	11	73.7%	21.4%
		부실(1)	4	3		

위 표는 Random Forest에 의한 예측결과를 나타낸 것이다. 위 표에서 0은 정상기업이고 1은 부실기업을 의미한다.

첫 번째 실험에서 재무데이터만 반영한 ‘Model 1’ 으로 예측했을 때 실제부실기업 중 부실기업을 맞게 예측한 Recall(재현율)은 14.3%이다. 그리고 전체 예측 값 중 정상기업을 정상기업으로 예측하고 부실기업을 부실기업으로 맞게 예측한 결과를 더한 비율인 Accuracy(정확도)는 63.2%이다.

두 번째 실험에서, 비재무데이터만 반영한 ‘Model 2’ 로 예측했을 때는 Recall(재현율)이 7.1%이고 Accuracy(정확도)는 75.4%이다.

세 번째로 재무데이터와 비재무데이터를 모두 반영한 ‘Model 3’ 으로 예측했을 때 Recall(재현율)은 21.4%이고 Accuracy(정확도)는 73.7%이다.

5.3.2 Logistic Regression 모형의 평가

[표 14] Logistic Regression 모형의 예측 결과

Logistic Regression			Actual		Accuracy	Recall (Sensitivity)
			정상(0)	부실(1)		
Model 1	Predicted	정상(0)	43	13	77.2%	7.1%
		부실(1)	0	1		
Model 2	Predicted	정상(0)	40	12	73.7%	14.3%
		부실(1)	3	2		
Model 3	Predicted	정상(0)	36	10	70.2%	28.6%
		부실(1)	7	4		

위 표는 Logistic Regression에 의한 예측결과를 나타낸 것이다. 첫 번째로, 재무데이터만 반영한 ‘Model 1’ 으로 예측했을 때 부실기업을 부실로 예측한 Recall(재현율)은 7.1%이다. 그리고 전체 데이터에서 정상기업을 정상기업으로 예측하고 부실기업을 부실기업으로 맞게 예측한 결과를 더한 비율인 Accuracy(정확도)는 77.2%이다.

두 번째로, 비재무데이터만 반영한 ‘Model 2’ 로 예측했을 때 Recall(재현율)은 14.3%이고 Accuracy(정확도)는 73.7%이다.

세 번째로 재무데이터와 비재무데이터를 모두 반영한 ‘Model 3’ 으로 예측했을 때 Recall(재현율)은 28.6%이고 Accuracy(정확도)는 70.2%이다.

평가지표의 예측결과가 우수하면 신용평가모형 개선에 반영하고자하였으나 아쉽게도 재현율(Recall)이 낮은 문제가 있었다.

제 6 장 소상공인 정책자금 부실관리

6.1 부실관리의 목적

소상공인 정책자금의 부실관리 목적은 정책자금의 지속가능한 지원으로 보다 많은 기업이 혜택을 받을 수 있도록 부실관리시스템을 체계화하여 국고재원의 손실을 줄이는 것이다. 부실관리는 기업의 부실징후에 조기 대응하여 기업의 재무구조를 개선하고 부실로 인한 손실을 최소화하도록 지원하며 기업부실에 의한 금융기관의 건전성 악화를 방지하고 국가경제를 안정화한다.

우리나라의 경우 IMF시기 이후 많은 부실기업이 발생하였고 소상공인의 폐업률은 상당히 높은 편이다. 아래 표를 보면 창업 3년 후 소상공인의 평균 생존률은 39.2%로 절반도 되지 않는 비율이다.

[표 15] 소상공인 생존율(평균)

(단위 :%)

창업연도	1년 후	2년 후	3년 후	4년 후	5년 후
2007년	61.8	49.3	41.5	35.8	30.9
2008년	60.9	48.1	39.7	33.4	29.0
2009년	60.1	46.9	38.0	32.2	27.3
2010년	60.0	46.3	38.2	31.9	27.5
2011년	59.8	47.3	38.8	32.8	-
2012년	60.1	47.5	39.1	-	-
2013년	62.4	49.5	-	-	-
2014년	62.7	-	-	-	-
평균	61.0	47.8	39.2	33.2	28.7

자료 : 소상공인시장진흥공단, 통계청 기업생멸행정통계

대출 이후 사후관리시스템이 제대로 작동하지 못하면 부실기업이 신속히 퇴출되지 못하거나 파산한 기업의 자산처리가 신속하게 이루어지지 못한다. IMF이후 부실기업의 처분지연에 따라 발생한 여러 사회적 비용을 계산하면 수조원에 이른다고 한다. 그러므로 부실을 관리하고 조기에 방하기 위한 시스템의 중요성이 매우 크다고 할 수 있다.

소상공인시장진흥공단의 소상공인 정책자금 직접대출 제도는 2016년부터 시행되었으므로 아직 축적된 데이터가 적고 부실관리 시스템이 다른 정책자금 제도에 비하여 미흡한 면이 있다. 그러므로 부실관리에 대한 연구의 필요성이 큰 상황이며, 부실관리 시스템의 보완을 위해 정책자금을 지원받고 상환중인 기업을 대상으로 하여 부실예측 연구를 진행하였다.

6.2 변수선정

변수는 5.1 부실예측에서 사용된 재무적변수와 비재무적 변수에 사후적 예측변수로 연체일수를 추가로 선정하였다.

6.3 실증분석

랜덤포레스트(Random Forest)와 로지스틱회귀(Logistic Regression)모형을 이용하여 자료를 분석하였으며 변수의 구성에 따라 각각 3가지 모형을 만들어 실험을 실시하였다. Model 1은 재무적 변수와 연체일수를 반영한 것이고 Model 2는 비재무적변수와 연체일수를 반영한 것이다. 그리고 Model 3은 Model 1과 Model 2에서 가장 기여도가 높은 상위 5개의 재무적변수와 비재무적변수, 그리고 연체일수를 반영한 것이다. 재무적 변수 5가지는 기여도순으로 영업이익(OpProfit), 당기순이익(NetProfit), 매출액

(Sales), 총자산회전율(AssetTurnover), 부채비율(DebtRatio)이며 비재무적 변수 5가지는 소호점수(SOHO), CB점수(CB), 업력(Life), 사업성점수(Busi), 종업원수(Worker)이다.

6.3.1 Random Forest 모형의 평가

[표 16] Random Forest 모형의 예측 결과

Random Forest			Actual		Accuracy	Recall (Sensitivity)
			정상(0)	부실(1)		
Model 1	Predicted	정상(0)	42	8	84.2%	42.9%
		부실(1)	1	6		
Model 2	Predicted	정상(0)	42	4	89.7%	71.4%
		부실(1)	2	10		
Model 3	Predicted	정상(0)	42	6	86.2%	57.1%
		부실(1)	2	8		

위 표는 Random Forest에 의한 예측결과를 나타낸 것이다. 재무데이터와 연체일수를 반영한 ‘Model 1’ 으로 예측했을 때 부실기업을 부실로 예측한 Recall은 42.9%이고 정상기업을 정상기업으로 예측하고 부실기업을 부실기업으로 예측한 Accuracy는 84.2%이다. 두 번째로, 비재무데이터와 연체일수만 반영한 ‘Model 2’ 로 예측한 결과 Recall은 71.4%이고 Accuracy는 89.7%이다. 세 번째로 비재무데이터, 연체일수를 모두 반영하여 ‘Model 3’ 으로 예측한 결과 Recall은 57.1%이고 Accuracy는 86.2%이다. 결론적으로 Random Forest에서는 ‘Model 2’ 가 가장 뛰어난 성능을 보였다.

6.3.2 Logistic Regression 모형의 평가

[표 17] Logistic Regression 모형의 예측 결과

Logistic Regression			Actual		Accuracy	Recall (Sensitivity)
			정상(0)	부실(1)		
Model 1	Predicted	정상(0)	39	2	89.5%	85.7%
		부실(1)	4	12		
Model 2	Predicted	정상(0)	41	4	89.5%	71.4%
		부실(1)	2	10		
Model 3	Predicted	정상(0)	40	1	93.0%	92.9%
		부실(1)	3	13		

위 표는 Logistic Regression에 의한 예측결과를 나타낸 것이다. 재무데이터와 연체일수를 반영한 ‘Model 1’ 으로 예측했을 때 부실기업을 부실로 예측한 Recall은 85.7%이고 정상기업을 정상기업으로 예측하고 부실기업을 부실기업으로 예측한 Accuracy는 89.5%이다. 그리고 비재무데이터와 연체일수만 반영한 ‘Model 2’ 로 예측한 결과 Recall은 71.4%이고 Accuracy는 89.5%이다. 그리고 재무데이터, 비재무데이터, 연체일수를 모두 반영한 ‘Model 3’ 으로 예측했을 때 Recall은 92.9%이고 Accuracy는 93%이며 가장 좋은 예측결과를 나타냈다.

6.4 부실관리 방안

소상공인 정책자금을 지원받은 기업의 부실로인한 피해를 최소화하기 위해서는 연체 10일 시점부터 사고 기업으로 관리할 필요성이 있다. 소상공인 정책자금 수혜기업을 대상으로 실증분석한 결과 Logistic Regression 모형에서 Accuracy 93%, Recall 92.9%로 부실예측이 가능하다는 성과를

얻을 수 있었기 때문이다. 현재 소상공인 정책자금 지원제도에 따르면 이자 연체 15일 시점부터 사고 기업으로 관리할 수 있으나 연체일 수 10일이 넘으면 부실이 거의 확실시된다는 것을 실험을 통해 알게 되었으므로 사고기업 관리기준을 10일로 변경하고 이 시점부터 리스크 관리 부서에서 적극적인 회수노력을 기울인다면 기금의 손실을 줄이는데 도움이 될 수 있을 것이다.

제 7 장 소상공인 정책자금 지원제도 개선 방안

본 연구를 통해 알아본 소상공인 정책자금 지원제도의 개선 방안과 지원기관의 역할은 다음과 같다.

첫째, 정책자금 지원기업의 부실예측과 관리를 위한 비재무적 평가지표를 개발해야 한다. 소상공인은 금융이력이 부족하고 재무적정보가 충분하지 않으므로 그 부족한 부분을 보완할 대안적 정보가 필요하다. 본 연구에서는 재무적 정보의 대안으로 기업의 비재무적정보를 활용하여 실험하였으며 비재무적정보가 부실 예측력을 제고할 수 있음을 검증하였다. 실험에 사용한 비재무적정보는 기업의 수많은 정보 중 일부분에 해당하는 정보였음에도 예측력을 높일 수 있었으며, 향후 더 많은 정보가 수집된다면 예측력은 더욱 향상될 수 있을 것으로 기대된다. 기업평가를 위한 비재무적 평가지표의 개발은 재무적 정보 위주의 평가방식에 의해 금융지원을 받지 못한 기업에게 지원을 받을 수 있는 기회를 제공할 수 있다. 소상공인 지원기관은 그동안 소상공인을 지원하며 오랜 기간 축적된 데이터와 노하우를 보유하고 있으므로 이를 잘 활용한다면 소상공인 평가에 효과적인 지표의 개발이 가능할 것이다

둘째, 여신심사모형이 개선되어야 한다. 기업평가를 위해 더 상세한 정보가 수집되어야 하며 심사모형이 정교화되어야 한다. 앞의 5장에서 여러 개의 모형을 만들어 실험하였으나 예측모형에 활용된 변수들의 값이 정밀하지 못하여 예측에 어려움이 있었다. 하지만 변수로 사용된 개별 평가항목들의 점수가 낮을수록 부실기업이 많이 분포하고 있었으므로 평가기능이 없는 것으로 보이지는 않는다. 실험에 사용된 변수 중 일부를 나타낸 아래 표를 보면, 기업의 여신심사과정에서 조사하는 업력,

SOHO점수, CB신용평점, 금융부채, 연매출액, 1인당 매출액의 수치가 낮은 구간에 부실기업이 많이 분포하는 것을 알 수 있다.

[표 18] 사고기업 분포비율

(단위 : %)

	업력	SOHO 점수	CB 신용평점	금융부채	연매출액	1인당 매출액
0%~25%	42	42	40	31	29	29
26%~50%	26	26	25	22	22	25
51%~75%	22	22	18	22	25	23
76%~100%	11	11	17	26	25	23
합계	100					

표를 보면 수치에 조금씩 차이가 있기는 하나 모든 항목에서 평가점수가 낮은 하위 그룹에 사고 기업이 많이 분포함을 알 수 있다. 업력, SOHO점수(CB신용평점+기업신용도), CB신용평점 등 각 변수로 사용된 항목들의 최고점을 100%로 환산하여 점수에 따라 4구간으로 나누어보면 각 항목들의 평가점수가 낮을수록 사고 기업이 가장 많이 분포하는 데이터는 '업력'과 'SOHO점수'이다. 그러므로 위 표에서 가장 유의미한 항목을 업력과 SOHO점수로 보고 그 중 하나인 SOHO점수의 최고점을 100%로 환산하여 수치를 비교해보았다. 아래 표를 보면 SOHO점수가 낮을수록 2금융권을 많이 이용하고 있음을 알 수 있다.

[표 19] 부실기업의 2금융권 이용 현황

(단위 : 건, 복수선택)

	부실기업의 SOHO점수				합계
	0%~25%	26%~50%	51%~75%	76%~100%	
중·고금리 대출 (10%이상)	27	21	9	1	58
카드론 (장기카드대출)	21	20	9	1	51
현금서비스 (단기카드대출)	10	3	2	0	15
저축은행 대출	7	2	1	1	11
보험회사 대출	7	8	9	2	26
지역농협,수협,신협 대출	11	13	8	1	33
새마을금고 대출	3	4	0	3	10
캐피탈 대출 (자동차,기계 제외)	25	6	2	4	37
자동차 대출	16	10	5	3	34
자동차 할부	25	32	28	21	106
자동차 리스	20	24	10	10	64
기계 리스	25	27	23	5	80
기타2금융할부	7	2	3	2	14
합계	204	172	109	54	539

2금융권을 많이 이용하는 다중채무자는 좋은 평가점수를 받을 수 없다. 위 표들은 신용평가를 구성하는 세부 항목들의 점수가 낮을수록 부실 확률이 높다는 것을 보여주고 있으므로 이 항목들이 평가에 도움을 주는 것은 맞으나 부실예측 연구에서 유의한 결과 값을 얻지 못한 것은 이 평가항목의 정보들이 더 자세하게 수집될 필요성이 있음을 보여준다. 예를 들면, 현행의 사업성평가모형은 한 가지 평가항목에 대하여 A, B, C, D의 4가지 선택지 중 하나를 선택하여 평가하게 되어있으나 선택지를 2배로 늘려서 A, B, C, D, E, F, G, H의 8가지 선택지 중 하나를 택

하도록 세분화한다면 이전보다 더 정밀한 데이터를 얻을 수 있을 것이고 유의미한 결과를 얻는데 도움이 될 것이다. 또한 심사모형 정교화를 위해 평가항목도 개선되어야 한다. 본 연구에서는 몇 가지의 정보들이 변수로 명시되어 있으나 연구를 위해 수집한 정보는 훨씬 많았다. 하지만 기업의 부실과 연관성이 없는 것으로 확인되어 실험에서 제외하였다. 정책자금 지원기관은 대출사례를 면밀히 분석하여 평가기능이 낮은 항목은 심사목록에서 제외하고 심사를 강화해야 할 부분에 집중해야 한다. 지난해 4월부터 연대보증제도가 폐지되었으므로 정책자금의 손실을 최소화하기 위해서는 심사절차를 강화할 필요가 있다. 하지만 모든 항목의 심사를 강화하면 심사에 소요되는 시간이 상당히 증가하므로 심사의 효율성을 제고하기 위해 평가항목을 분석하여 재구성할 필요가 있다.

셋째, 정부 관련 부처 간 융합 연구를 수행해야 한다. 소상공인대출의 가장 큰 문제점은 정보가 부족하다는 점이다. 기업 정보의 부재로 인하여 기업을 제대로 평가하기가 어렵고 관리도 쉽지 않다. 하지만 빅데이터(Big Data), 인공지능(Artificial Intelligence), 사물 인터넷(Internet of Things) 등 정보통신기술(Information and Communication Technology)의 발달에 의하여 큰 변화가 예상된다. 비재무적 정보만으로도 여신심사가 가능한 시스템이 개발되었고, 외국과 일부 핀테크 기업에서는 이미 대출서비스를 제공하고 있다. 우리나라에서는 개인정보보호법, 정보통신망법, 신용정보법이 개정되어야 본격적인 서비스가 가능할 것이다. 규제 문제가 해결된다면 앞으로 개인의 데이터 수집과 관리가 용이해지고 비재무적 데이터를 대량 수집하는 것이 가능해진다. 소상공인은 우리나라 전체 기업 수의 85.3%를 차지할 정도로 많으며 대출규모가 크지 않으므로 소수의 기업이 많은 자금을 조달하는 대기업에 비하여 규모대비 건당 취급비용이 높는데, 정보통신기술의 발달로 컴퓨터시스템이 대부분의 역

할을 수행하게 된다면 비용문제도 해결될 수 있을 것이다. 관련부처 간 시스템이 구축된다면 정보를 통합 관리하여 정책자금 중복지원문제도 개선될 수 있을 것이다. 그리고 여신심사에 필요한 정보도 수월하게 수집할 수 있게 되어 서류발급을 위해 여러 기관을 방문하느라 영업을 하지 못하여 발생하는 매출 손실이나 준비하는 수고를 덜어줄 수 있다. 소상공인 지원기관은 여신심사와 관리를 위한 지표와 예측모형의 개발방향을 설정하고 관련부처의 협조를 받아 기능 구현을 위한 시스템을 공동 개발한다면 자금조달에 어려움을 겪는 소상공인을 위한 대출시스템이 빠른 시일 내에 구축될 수 있을 것이다.

넷째, 소상공인 정책자금 지원대상자를 위한 교육과정을 마련해야한다. 기업의 부실을 잘 예측하기 위해서는 기업의 정보가 신뢰성이 있어야 한다. 재무제표와 기업의 기타 재무정보들이 허위로 작성되지 않도록 윤리교육과 의식개선교육이 지원되어야 할 것이며 금융교육과정을 마련하여 대표자의 역량을 강화하고 주체적으로 기업의 재무정보를 잘 관리할 수 있도록 해야 한다. 남윤형(2016)의 연구에서는 다수의 소상공인들이 폐업 후 재도전하는 회전문 창업을 선택하고 있으며, 제조업 소상공인의 경우 폐업 후 다시 제조업으로 재창업하는 비중이 49.8%로 가장 높은 것으로 나타났다. 폐업과 창업의 과정에서 발생하는 비용에 의해 재창업 후 경제적 여유가 없어지게 되고 자산이 소진되면 융자지원을 필요로 하게 된다. 이러한 과정이 반복되면 기업의 부실로 이어질 수 있고 많은 사회적 비용이 발생된다. 그러므로 교육을 통해 소상공인의 역량을 강화하여 창업과 폐업의 악순환을 줄이고 사업의 영속성을 제고할 수 있는 방안을 마련하는 것이 바람직 할 것이다.

다섯째, 소상공인 정책자금 지원제도의 홍보가 필요하다. 서민금융연구

원이 2019년 1월 발표한 대부업·사금융시장이용자 및 업계동향조사 분석 자료에 의하면 대부업체와 사금융을 모두 이용하고 있는 자영업자의 비중이 높은 것으로 조사되었다. 금리가 높은 대부업체를 이용하는 이유는 아래 표를 보면 어디서 자금을 조달할지에 대한 정보가 부족하여 고금리 대출을 이용하는 비중이 23%나 차지하는 것을 알 수 있다. 여신심사과정에서 기업의 재무적상태가 우수하여 1금융권 저금리대출을 이용할 수 있는데도 2금융 고금리대출을 이용하는 기업이 있었는데, 대표자와 면담을 해보면 실제로 정보부족의 원인이 컸다.

[표 20] 대부업·사금융시장이용자 설문조사

대부업체 이용 목적	응답비중(%)
금융기관에서 필요할 만큼 빌릴 수 없어서	63.5
빨리 대출해 줘서	26.0
어디서 돈을 빌려야 될지 모르는 상황에서 대부광고 및 전화, 문자 등을 보고	23.2
서류가 간단해서	15.7
대부업의 비밀보장성 때문	8.0
기타	8.2

자료 : 서민금융연구원, 대부업·사금융시장이용자 및 업계동향조사분석(2019)

여섯째, 정책자금의 목적을 고려하여 사회적 배려가 필요한 기업이 지원을 받을 수 있도록 정책자금 지원제도를 개선해야 한다. 신용등급이 상위등급이고 재무적 여건이 우수한 기업은 정책자금 뿐만 아니라 민간 금융기관에서도 자금을 지원을 받을 수 있다. 그러므로 모든 기업에게 동일한 이율과 상환조건이 적용되는 대출제도를 개선하여 사회적으로 배려가 필요한 기업이 많은 지원을 받을 수 있도록 할 필요가 있다. 현재 소상공인을 대상으로 하는 정책자금 지원제도를 살펴보면 은행보다 낮은 저금리대출이 대부분이고 이러한 자금은 조건이 까다로워 정작 지원받아

야할 기업이 지원을 받지 못하는 경우가 많다. 심사에 탈락된 기업은 자금을 조달하지 못하여 제2금융권이나 사금융시장의 고금리를 이용하게 되고, 높은 이자 부담으로 상황이 더욱 악화된다. 그러므로 정책자금의 목적을 유념하여 심사조건을 차별화한 평가모형을 마련할 필요가 있다. 그리고 고금리상품을 이용하는 소상공인을 대상으로 금리를 차등 적용하는 중금리대 대출을 정부가 적극적으로 지원한다면 민간의 고금리 시장도 경쟁에 의해 점차 금리가 낮아질 수 있을 것이고 신용이 낮은 기업의 선택폭이 넓어지며 이자 부담을 줄일 수 있을 것이다.

제 8 장 결론

본 연구는 조선기자재 분야 소상공인 정책자금의 부실을 예측과 관리 측면에서 살펴보고 정책자금 지원제도의 개선방안을 제시하였다. 표본 245개 기업을 랜덤포레스트(Random Forest)와 로지스틱회귀분석(Logistic Regression)을 이용하여 연구한 결과 대출 전 여신심사단계에서 부실기업을 예측하는 실험에서는 유의성을 검증하기 어려웠으나 대출 후 부실관리를 위한 실험에서는 로지스틱회귀모형에서 높은 예측 결과를 얻을 수 있었다. 그리고 재무적 변수들로만 실험한 모델보다 재무적 변수와 비재무적 변수를 혼합하여 실험한 모델이 더 높은 예측확률을 보임으로써 비재무적 지표 개발의 중요성과 향후 연구방향에 대하여 생각해 볼 수 있었다.

본 연구는 제조기반기술분야 경쟁력 강화를 위한 소상공인 정책자금의 수혜기업 데이터를 실증분석한 최초의 연구라는 점에서 의의가 있다. 그동안 많은 예산이 소상공인 정책자금 목적으로 투입되었고 직접대출의 규모가 매년 증가하고 있으므로 소상공인 직접대출의 효과와 개선방안 등이 검토되어야 함에도 아직 선행연구가 없었다. 신용평가모형이나 부실관리시스템의 검증과 개선에는 많은 시간이 필요한데 본연구로 인하여 향후 연구의 시행착오에 걸리는 시간을 줄이고 소상공인의 발전을 위한 정책자금 지원제도의 개선에 기여할 수 있기를 바란다.

하지만 본 연구는 다음과 같은 한계점을 가지고 있다.

첫째, 다양한 연구방법론이 적용되지 못하였다. 예측력이 우수한 소상공인 정책자금 부실예측 모형 개발을 위해 보다 다양한 연구방법론을 적

용해될 필요가 있다. 본 연구에서는 랜덤포레스트와 로지스틱 회귀분석을 이용하여 분석하였으나 부실예측에 적용할 수 있는 연구방법론이 다양하므로 여러 가지 시도를 통해 예측력을 더 높일 수 있는 모형을 개발할 수 있을 것이다. 본 연구에서 다루지 않은 새로운 머신러닝 기법이나 통계적 모형을 적용하여 다양한 모형을 개발하고 모형 간 성능을 분석하는 과정을 거쳐 소상공인 정책자금에 최적화된 부실예측모형을 개발한다면 부실률을 줄이는데 기여할 수 있을 것이다.

둘째, 연구의 대상이 제한적이었다. 이번 연구는 조선기자재 산업에 초점을 두었지만 다른 업종에 대한 연구도 진행하여 업종별 차이점이 있는지 비교분석해볼 필요가 있다. 업종별 차이점이 발견된다면 향후 여신심사와 관리 시 업종별 특성을 고려하는 방향으로 제도가 개선되어야 할 것이다. 특히 조선기자재 분야와 같은 경기민감 업종과 그렇지 않은 업종간의 차이점이 발견된다면 자금지원 규모와 사후관리 계획을 조정하여 위험성이 높은 업종이 더 면밀히 관리되도록 해야 할 것이다.

참고문헌

강종만, 홍성희, 1999. *부실예측모형의 적합성 분석*, 증권금융연구 5(1), pp. 83-110.

강치형, 신해수, 2015. *회원제 골프장 기업의 부도 예측 모형개발*, 관광연구논총 27(4), pp.241-269.

권영린, 2002. *호텔기업의 부실예측모형에 관한 연구*, 호텔관광연구 9, pp.55-74.

김상봉, 조경준, 2011. *부도예측모형을 이용한 기업부실화의 원인분석*, 시장경제연구 40(1), pp.85-106.

김시중, 2016. *부실예측모형의 예측력 비교*, 한국엔터테인먼트산업학회논문지, 10(4), pp.343-350.

김용덕, 이인자, 2002. *국내정보통신산업의 부실분석 및 예측모형에 관한 연구*, 정보통신정책연구 9(1), pp.163-198.

남기정, 이동명, 진로, 2019. *비재무정보를 이용한 창업기업의 부실요인에 관한 실증연구*, 벤처창업연구, 14(1), pp.139-149.

남윤형, 2016. *소상공인 회전문창업 실태와 해법의 실마리*, KOSBI 중소기업 포커스 16(3).

문종건, 황보운, 2014. *횡령·배임 및 최대주주변경을 고려한 부실기업예측모형 연구*, 벤처창업연구 9(1), pp.119-132.

문종건, 황보운, 2014. *횡령·배임 및 최대주주변경을 고려한 부실기업예측모형 연구*, 벤처창업연구 9(1), pp.119-132.

민재형, 정철우, 2007. *기업도산예측을 위한 이진분류기법의 개발*, 한국경영정보학회 학술발표대회 논문집, pp.619-624.

박정식, 신동령, 2016. *경영분석*. 다산출판사.

사)서민금융연구원, 2019. *대부업·사금융시장이용자 및 업계동향조사분석*, 서울:사)서민금융연구원.

신용보증재단중앙회, 2018. *2018 소상공인 금융실태조사 보고서*, 대전:신용보증재단중앙회.

유승규, 박정로, 최재규, 김재준, 2009. *다변량관별분석(MDA)기법을 이용한 외부감사 건설 기업의 부도 예측 모형개발에 관한 연구*, 대한건축학회 학술발표대회논문집 29(1), pp.653-656.

윤일현, 2011. *금융회사의 도산예측모형에 관한 연구*, 경영경제연구 34(1), pp.1-17.

윤종식, 권영식, 2007. *SVM을 이용한 소상공인 부실예측모형*, 대한산업공

학회 춘계공동학술대회논문집, pp.826-833.

이수현, 박정민, 이형용, 2015. 데이터마이닝 기법을 활용한 비외감기업의 부실화 유형 분석, 지능정보연구 21(4), pp.111-131.

이영찬, 2004. 인공신경망과 Support Vector Machine의 기업부도예측 성과 비교-Support Vector Machine의 유용성을 중심으로-, 한국지능정보시스템학회 학술대회논문집, pp.211-218.

이주희, 동학림, 2018. 소상공인의 자금공급 확대를 위한 빅데이터 활용 방안연구, 벤처창업연구, 13(3), 125-140.

이훈영, 2012. 연구조사방법론. 청람.

장영재, 김현중, 조형준, 2016. 데이터마이닝. 한국방송통신대학교출판문화원.

전현우, 정용화, 신동휴, 2011. 상장폐지기업의 부실예측모형에 관한 연구-거래소시장을 중심으로-, 국제회계연구, pp.331-362.

표수진 등, 2018. 머신러닝 기반 저축은행의 부실 징후 예측 및 업권 위험지표 지수 개발, 한국정보과학회, pp.835-837.

허광복, 2018. 비외감 중소기업의 이익조정에 관한 연구-금융기관 제출 재무제표의 인증 수준별 비교-, 세무와 회계저널 19(3), pp.185-213.

Altman E. I., 1968. *Financial Ratios, Discriminant Analysis and the Prediction Corporate Bankruptcy*, Journal of Finance Vol.23, pp.589-609.

Beaver, W. H., 1966. *Financial Ratios as Predictors of Failure*, Journal of Accounting Research, pp.71-111.

Davis, J. & Goadrich, M., 2006. *The relationship between Precision-Recall and ROC curves*, In Proceedings of the 23rd international conference on Machine learning, pp.233-240.

L. Breiman, 2001. *Random forests, Machine Learning*, Kluwer Academic Publishers 45(1), pp.5-32.

Ohlson, J. S., 1980. *Financial Ratios and the Probabilistic Prediction of Bankruptcy*, Journal of Accounting Research, pp.109-131.

[Online Available Only] Available at:

<https://ko.wikipedia.org/wiki/%EB%9E%9C%EB%8D%A4%ED%8F%AC%EB%A0%88%EC%8A%A4%ED%8A%B8> [Accessed 1 April 2019]

[Online Available Only] Available at:

https://ko.wikipedia.org/wiki/%EB%A1%9C%EC%A7%80%EC%8A%A4%ED%8B%B1_%ED%9A%8C%EA%B7%80 [Accessed 1 April 2019]

감사의 글

벌써 석사과정 졸업을 앞두고 있다는 사실에 감회가 새롭습니다. 대학교 졸업 후 다시 학생으로 돌아와 많은 것을 배우고 전문의 폭을 넓힐 수 있는 기회를 가질 수 있게 되어 감사했고 정말 뜻깊은 시간을 보낼 수 있었습니다. 석사 학위 논문을 준비하는 과정에서 크고 작은 어려움이 있었지만 많은 분들이 도움을 주셨기에 포기하지 않고 최선을 다할 수 있었습니다. 부족함이 많은 저에게 아낌없는 응원과 격려를 보내주신 분들에게 이 기회를 빌어서 감사의 인사를 전하고 싶습니다.

먼저 부족한 제 연구를 지속적인 관심으로 이끌어 주시고 세심하게 지도 해주신 이재민 지도교수님께 존경과 감사의 말씀을 드립니다. 그리고 바쁘신 와중에도 심사를 맡아 주시고 포기하지 않도록 격려해주신 이기환 교수님께도 진심으로 감사드립니다. 그리고 논문의 주제 선정부터 정말 많은 도움을 주신 임상섭 교수님께 진심으로 감사드립니다. 이 논문의 완성은 소중한 시간을 아끼지 않으시고 도와주셨던 교수님이 계셔서 가능했습니다. 그리고 좋은 강의로 많은 가르침을 전해주신 모든 교수님들께 진심으로 감사드립니다. 배움의 과정에서 연구를 시작할 수 있는 계기와 기회를 만들어주셔서 감사합니다. 앞으로 더 노력하며 성장해나가겠습니다. 그리고 논문을 쓰는 과정에서 많은 격려를 보내준 대학원 동기, 선후배님들에게도 감사의 마음을 전합니다.

직장을 다니며 논문을 준비하는 것은 결코 쉬운 일이 아니었지만 많은 분들의 도움이 있었기에 가능했습니다. 논문을 완성하는데 큰 힘이 되어주시고 격려해주신 박규식 심사역님, 많은 조언을 해주신 김삼덕 심사역님 그리고 학업을 응원해주신 소상공인시장진흥공단 서근하 본부장님을 비롯한 회사 선후배님들에게 감사의 말씀을 전하고 싶습니다. 그리고 지면을 통해 미처 언급하지 못했지만 그동안 도움주셨던 많은 분들께 진심으로 감사하다는 말씀을 전합니다.

끝으로 언제나 저를 믿고 아낌없는 사랑으로 응원해주시는 존경하는 아버지, 어머니 그리고 언니에게 감사의 마음을 전합니다.