

머신러닝을 활용한 지방정부 재정건전성 영향요인 탐색

장 유 미*
유 한 별**

국문요약

본 연구는 지방자치단체의 재정건전성에 미치는 영향을 확인하기 위해 인구구조와 사회·경제적 환경 요인에 초점을 맞추어 분석을 진행하였다. 분석 결과, 사회·경제적 환경 요인이 지방자치단체 재정건전성에 주요한 영향을 미치는 것을 확인하였다. 구체적으로, 복지 분야와 경제 분야가 영향을 미쳤으며, 복지 분야의 경우, 사회복지 예산의 비중이 채무상환비율과 예산대비채무비율에 음의 영향을 미치는 것을 확인하였다. 경제 분야의 경우, 머신러닝을 통한 분석 결과 산업별 영향력을 확인하였으며, 로지스틱 회귀분석에 의해 계수를 확인한 결과 통계적으로 유의미하진 않지만 농업 및 어업 등의 1차 산업 분야는 통합재정수지에 양의 영향을 미치는 반면, 서비스업 분야는 음의 영향을 미치는 것으로 나타났다. 분석 결과를 통해 연구는 지방자치단체의 중앙정부의 이전재원에 대한 의존을 확인하였으며, 장기적으로는 지방자치단체의 지방세의 탄력적 운영을, 단기적으로는 지방세수 확대를 위한 지방자치단체 경제적 기반 활성화를 제안하였다.

주제어: 지방정부, 지방재정, 재정건전성, 지방소멸, 머신러닝

I. 서론

지방자치제도는 지역의 균형발전과 민주주의의 실현을 위한 가장 중요한 수단 중 하나이다. 우리나라에서는 1995년 본격적으로 지방자치가 실시되었으며, 지방정부는 분권화를 통해 자율적으로 지방정부가 원하는 정책을 형성 및 집행할 수 있게 되었다. 지방자치를 시행함으로써 가장 큰 변화는 재정사무에 관한 권한을 분권화하여 지방정부에 자율성과 권력을 갖도록 보장해 준다는 것이다. 이로 인해 지방정부는 주민들이 원하는 재화와 서비스를 효율적으로 제공할 수 있게 되었다. 지방재정 운영이란 지방자치단체가 내·외부적 자원으로부터 재원을 조달하고 이들 재원을 주민복지와 지역개발을 위한 다양한 재정수요에 지출하는 일련의 과정이라고 할 수 있다(김상길, 2021). 따라서 지방정부의 사업을 추진할 수 있는 원동력인 재정을 효율적이고 효과적으로 관리하는 것은 지방정부에게 있어서 매우 중요한 역할이라고 할 수 있다.

* 제1저자

** 교신저자

지방자치단체의 재정을 대상으로 한 연구는 주로 정부의 운영 방향을 확인할 수 있으며, 정책 산출의 지표로서 타당성이 대체로 인정되는 지방재정지출을 대상으로 한 연구가 활발히 이루어져 왔다(정연택·이명숙, 2007). 한편, 지방재정 건전성과 관련한 연구 또한 진행되어 왔으며, 주로 지방재정건전성 확보를 위한 제도적 개선방안을 중심으로 이루어져 왔다. 반면, 지방자치단체의 재정건전성에 영향을 미치는 요인에 관하여 실증 분석한 연구는 비교적 적은 실정이다. 지방정부의 재정건전성은 지방정부가 시민들에게 양질의 서비스를 제공할 수 있는 능력을 나타내며, 지방자치단체의 재정적 능력이 악화될 경우, 이로 인해 발생하는 지역의 문제들은 시민들에게 직접적으로 영향을 미칠 수 있기 때문에 지방정부의 재정건전성 관리는 매우 중요하다. 특히, 우리나라는 급속도로 악화되고 있는 저출산·고령화 현상과 이에 더하여 코로나19 팬데믹 등의 문제로 인한 세입의 감소와 지출의 폭발적 증가, 복지비 지출의 증가 등으로 재정위기를 겪고 있다. 이러한 상황에서 지방재정건전성에 영향을 미치는 요인에 대한 연구는 시의적절하다고 볼 수 있다.¹⁾

본 연구의 목적은 지방자치단체의 재정건전성 확보를 위해 지방재정건전성에 영향을 미치는 요인을 확인하는 것이다. 특히, 연구는 지방자치단체의 환경적 요인을 중심으로 분석을 진행한다. 지방자치단체의 환경적 변화는 지역의 경제발전과 세입 확보에 큰 영향을 미치는 요인 중 하나이기 때문에 환경이 지방재정의 건전성에 미치는 영향을 확인하는 것은 매우 중요하다. 그럼에도 불구하고 대부분의 지방재정건전성 관련 연구들은 주로 지방자치단체의 재정·재무적 요인이 재정건전성에 미치는 영향에 중점을 두어 연구를 진행하였다. 이에 본 연구는 지방정부의 인구구조의 변화와 사회·경제적 요인 등 환경적 요인에 초점을 맞추어 지방재정건전성 영향요인을 확인하고 향후 지방재정건전성 확보를 위한 시사점을 제시하고자 한다.

II. 이론적 논의

1. 재정건전성의 정의

재정을 학문적 개념으로 살펴보면 국가가 그 임무를 달성하기 위해 필요한 자금을 조달하고 사용하는 것을 의미한다(Mayer, 1981; 이진수 외, 2020). 일반적으로 공공분야에서 재정이란 정부가 공공수요를 충족하기 위해 필요한 재원을 조달하고 재산을 관리·사용하는 행위 또는 공공단체가 그 존립을 유지하고 활동하는데 필요한 재화를 취득·관리·사용하는 모든 행위를 의미한다(권영성 외, 2017; 김난진·김연태, 2011). 한편, 건전성의 개념은 재정의 개념보다 모호한 의미를 갖기 때문에 이에 대한 논의는 다양하게 나타내며, 그렇기 때문에 재정건전성과 관련하여 학자마다 다양한 정의를 제시하고 있다.

우리나라의 「지방재정법」 측면에서 보면, 동법 제3조에서 지방재정 건전성을 “지방자치단체는

1) 지방정부의 통합재정수지비율은 17개 광역자치단체 기준 2012년 4분기 0.1%에서 2021년 4분기 -2.9%로 급속히 악화되고 있다.

주민의 복리 증진을 위하여 재정을 건전하고 효율적으로 운영하여야 하며 국가의 정책에 반하거나 국가 또는 다른 지방자치단체의 재정에 부당한 영향을 미치게 하여서는 아니 된다.”고 규정하고 있다.

학문적 측면에서 재정건전성은 재정건강도, 재정지속가능성, 재정압박, 재정의무이행, 생산성, 성장성, 지출의 효율성, 재원배분의 합리성 등의 여러 가지 형태로 정의되어 왔다(임성일·서정섭, 2004). 재정건전성과 관련한 관점은 크게 정부의 재정적 의무와 관련한 미시적인 관점과 재정건전성 확보를 통한 정부의 시민에 대한 책임에 초점을 맞추는 거시적 관점 두 가지로 나눌 수 있다. 먼저, 미시적 관점의 경우는 단순히 정부의 재무적 측면과 관련하여 재정건전성을 정의한다. Berne & Schramm(1986)는 재정건전성을 “정부가 공공서비스를 수입증대를 통해 제공하며, 정부의 채무를 단기적 또는 장기적인 관점에서의 상환할 수 있는 능력”이라고 정의한다. 이와 비슷한 맥락으로 Ward and Dadayan (2009: 455)은 “장기적인 관점에서 세입과 세출의 균형을 맞추는 정부의 능력”이라고 정의한다(안국찬, 2013). 김동현 외(2013)는 재무적 관점에서 재정건전성은 일반적으로 정부의 채무상환능력이라고 할 수 있으며, 단기에 만기가 도래하는 국채원리금의 상황에 있어 용이성, 즉 유동성 확보 능력을 의미하는 안정성과, 장기적 관점에서의 채무상환능력을 의미하는 지속가능성을 포괄한다고 본다(라영재 외, 2012). 즉 재정건전성은 재정에서 수입과 지출의 균형점 등의 수지균형과 밀접한 관련이 있다고 보았다(김동현 외, 2013). 한편, 정부의 시민에 대한 책무와 관련한 거시적 관점도 존재한다. Ladd & Yinger(1989)는 재정건전성을 주민에게 공공서비스를 제공하는 지방자치단체의 능력이라고 정의하였다. 즉, 정부에서 계획한 예산과는 관계 없이 주민들에게 공공서비스를 제공하는 도시의 기본적·구조적 능력을 의미한다. Honadle *et al.*(2003)과 OECD(2013)는 재정건전성을 정부가 시민(유권자)들에게 지속적으로 적절한 서비스를 제공할 수 있도록 공공재정을 유지할 수 있는 능력이라고 정의하고 있다.

2. 재정건전성에 영향을 미치는 요인과 기계학습(machine learning)

지방재정건전성에 영향을 미치는 요인은 매우 다양하며, 크게 재정적 요인과 사회·경제적 요인으로 나눌 수 있다. 먼저 재정적 요인은 지방정부의 수입·지출 구조, 재정정책 등의 재정적 특성이 지방재정 건전성에 영향을 미친다고 본다. 재정적 요인은 주로 재정건전성에 대한 미시적 정의에서 기인된다. 즉, 재정건전성의 정의 그대로 정부의 세입과 세출, 채무상환능력, 재정수지 등이 재정건전성에 영향을 미친다고 보는 것이다. 우리나라에서는 지방재정 건전성에 영향을 미치는 요인을 탐색한 연구들이 다수 있었으며, 주로 지방정부의 세입과 세출, 지방채무, 지방채상환비율, 등을 영향요인으로 보고 지방재정건전성에 영향을 미치는 요인을 분석하였다.

한편, 사회·경제적 요인과 관련한 논의는 지방정부를 둘러싼 사회·경제적 요인들이 재정수요와 수입에 영향을 미쳐 지방정부의 재정형태에 영향을 미치고, 이는 궁극적으로 지방재정건전성에 영향을 미친다고 보는 관점이다. Honadle *et al.*(2003)은 지방정부의 재정건전성에 영향을 미치는 사회·경제적 요인은 매우 다양하지만, 지방정부의 사회·경제적 환경과 인구변화가 지역의 재정건

전성에 큰 요인으로 작용할 수 있다고 보았다. 즉, 이러한 요인들은 지방정부에 주거하는 개인들의 소득과 그들의 지방정부 서비스에 대한 서비스 수요에 영향을 미치며, 이는 곧 지방정부의 세입과 세출에 영향을 미치기 때문에 지방정부의 재정건전성에 영향을 미칠 수 있다는 것이다.

먼저, 장기 요양을 필요로 하는 노인층, 유·청소년, 출생률 등 다양한 노동력을 포함하고 인구변화는 지방재정 건전성에 핵심적인 요인으로 작용할 수 있다(Honadle *et al.*, 2003). Honadle *et al.*(2003)에 따르면, 인구변화는 서비스의 수요와 지불 능력에 영향을 미칠 수 있는데 노인층이 증가할 경우, 장기적인 복지정책에 대한 수요가 높아져 지방정부의 의무지출 및 복지비 지출이 증가하는 반면, 노동력은 감소하여 세입은 줄어들게 되어 지방재정 건전성에 영향을 미칠 수 있다.

우리나라의 경우, 인구구조가 급격히 변화하고 있으며, 이에 따라 모든 사회·경제적 부문에서 인구구조 변화에 따른 영향을 받고 있다. 인구 감소로 인한 충격은 지방을 중심으로 더 빠르고 광범위하게 초래될 가능성이 높다. 국가 전체적인 인구감소와 그에 따른 사회경제적 활력 쇠퇴로 인해 지방의 인구와 기업들이 좀 더 생존 가능성이 큰 수도권으로 쏠리는 현상이 가속화함으로써 지방소멸 가능성이 더욱 커지고 있으며, 이에 따른 지방의 세수 감소 및 복지비 지출의 증가가 나타나고 있다. 홍영교·최남희(2021)에 따르면, 실제로 인구구조의 변화가 지역의 경제활동인구 구조에 영향을 미쳐 고령인구가 높은 지역의 경우, 경제활동 인구 감소에 따른 공급 감소 및 성장 잠재력 약화 가능성을 제기하였다. 그들의 연구는 이로 인해 인구구조의 변화는 경제부문과 지역불균형에 영향을 미치며, 결국 지방재정 건전성에 부정적 영향을 미칠 수 있다고 지적하였다.

다음으로, 경제·사회적 환경의 경우, 그 지역이 갖는 인프라 및 시설과 지역경제 기반구성에 따라 영향을 받을 수 있다. 지역의 인프라 시설과 지역경제 기반구성은 지역 인구의 소득과 지역경제 성장뿐만 아니라 정부의 세출에도 영향을 미칠 수 있다(Honadle *et al.*, 2003). 지역 인프라와 지역경제 기반구성은 정부의 정책 방향을 결정하는데 중요한 역할을 한다. 따라서 지방정부의 지출에 영향을 미치며, 궁극적으로는 세입에 영향을 미칠 수 있다.

사회·경제적 요인을 파악하는 것은 재정적 요인에서 확인할 수 없는 지방정부의 재정건전성에 미치는 영향요인에 대한 총체적인 영향의 정도를 파악할 수 있으며, 이를 통해 근본적인 재정건전성 문제를 해결할 수 있다는 장점을 가진다. 이에 본 연구는 인구통계학적 변화와 사회·경제적 환경에 관한 변수를 수집하여, 기계학습(machine learning) 방법을 이용하여 이전에 진행되었던 통계적 방법에 대한 보완적이고 탐색적인 연구를 진행하고자 한다.

기계학습은 자료를 훈련자료(train set)와 시험자료(test set)로 나눈 후 훈련자료를 통해 모형을 구축한다. 이후 시험자료의 독립변수만을 해당 모형에 반영하여 종속변수를 예측하며, 이 값을 시험자료의 실제 종속변수 값과 비교를 통해 모형을 평가한다. 따라서 회귀계수의 정확한 추정보다 시험자료의 독립변수에 의한 종속변수 예측값과 실제 시험자료의 종속변수가 얼마나 일치하느냐에 목적이 있으며, 이를 통해 모형의 성능을 판단한다. 결국, 다중공선성(multi-correlation)을 고려하여 소수의 변수만을 반영하여 회귀계수를 추정하였던 기존의 회귀분석 보다 분석에 반영할 수 있는 변수를 좀 더 자유롭게 선택할 수 있다. 본 연구에서는 이러한 장점에 주목하여 다중공선성에서 비교적 자유로우며, 다양한 변수를 모형에 반영하여 모형을 추정할 수 있는 기계학습방법을

사용하여 지방재정건전성에 미치는 영향을 확인하고자 한다.

3. 지방재정건전성 영향요인 관련 선행연구 분석

우리나라 지방재정 건전성에 영향을 미치는 요인을 탐색한 연구들은 주로 지방정부의 세입과 세출, 지방채무, 재정건전성, 의회의 예산권한 등 재정적 요인을 주요 영향요인으로 보고 분석을 진행하였다. 먼저, 강남호(2011)의 연구는 지방정부의 재정건전성을 확인하기 위해 지방재정 건전성을 실증 분석하였다. 분석결과 지방재정의 자주재정력이 부족하며, 지방재정의 효율적 운영, 세입구조의 개선, 정책적 독립성 확보, 지방재정관리제도의 합리화 등이 필요하다고 제안하였다. 오영호·편도창(2012)의 연구는 지방자치단체의 재정특성이 지방채무 수준에 미치는 영향을 확인하기 위해 연구를 진행하였다. 분석결과, 재정자립도는 부채비율에 정적 영향을 미친 반면, 1인당 부채규모는 음의 영향을 미친 것을 확인하였다.

안국찬(2013)은 우리나라 지방자치단체의 건전성 문제를 다양한 지표를 통해 분석한 후 향후 건전성 제고방안을 제시하였다. 연구는 세출입현황, 재정자립도, 지방채무 잔액, 채무비율, 사회복지비 비중 등의 지표가 건전성에 미치는 영향을 분석하였으며, 분석결과 지방정부의 재정자립이 필요하고, 지방세입 감소 원인의 제고, 지방재정분석과 지방정부 재정운용의 효율화 및 감시체계를 확립해야 한다고 주장하였다. 위정복(2015)의 연구는 지방재정 건전성에 영향을 미치는 요인을 지방재정 수입·지출의 규모로 보고 분석을 진행하였다. 분석결과, 자치단체의 재정자립수준이 높아질수록 기초자치단체에서 재정수지가 악화될 반면 채무부담은 감소하였으며, 1인당 투자지출 증가는 재정수지에 부적 영향을 미치는 것을 확인하였다.

이성근 외(2016)는 지방정부의 재정건전성에 대한 논의를 세입·세출구조와 같은 재정구조적인 문제로 보고 지방자치단체의 세입에 대한 역량 및 세출구조가 재정건전성에 미치는 영향을 확인하였다. 분석결과 예산대비 채무비율에 영향을 미치는 요인은 재정자주도가 음의 영향을, 복지비 비율이 양의 영향을 미치는 것으로 나타났다. 이를 통해 지방자치단체에 재정운영상의 자율성을 부여해주면 예산대비 채무비율이 낮아지게 되어 재정건전성을 확보하는데 기여할 수 있다는 것을 알 수 있다. 이에 연구는 지방정부의 자주재원이 확충될 수 있도록 지방정부의 전권적 자율성을 인정할 필요가 있다고 주장하였다.

최예나·김상현(2017)의 연구는 지방재정 위기 문제 극복과 재정건전성 향상을 위해 지방재정 건전성에 영향을 미치는 지역요인을 탐색하였다. 이를 위해 통합재정수지비율과 부채비율에 영향을 미칠 수 있는 재정적 요인, 정책적 요인, 인구·사회적 요인 및 경제적 요인으로 나눠서 분석을 진행하였다. 분석결과, 지방정부의 재정자립도와 복지지출 비중이 낮고 지역개발지출 비중과 노령인구 비중이 높으면서, 인구 1,000명당 산업종사자 수가 많은 경우 지방재정 건전성이 높아진다는 것을 확인하였다(최예나·김상현, 2017).

이현정·조정래(2019)의 연구는 지방세 비과세·감면을 중심으로 지방자치단체 재정건전성에 영향을 미치는 요인을 분석하였다. 분석결과 지방세 비과세·감면이 단기적으로는 재정건전성에 부

정적인 영향을 미쳤지만, 장기적 관점에서는 유의미한 영향을 미치지 않는 것으로 확인되었다. 해당 연구는 분석결과를 바탕으로 지방세 비과세·감면의 사전 영향평가, 세입감소분에 대한 충분한 재정적 보전 등의 제도가 활성화 되어야 한다고 주장한다.

강경표·김준기(2020)의 연구는 재정민주주의가 재정건전성에 미치는 영향을 확인하기 위해 의회의 예산권한이 재정건전성에 미치는 영향을 실증 분석하였다. 분석결과 의회예산권한지수가 재정건전성에 부정적 영향을 미치는 것을 확인하였다. 분석결과를 바탕으로 연구는 재정준칙의 도입, 실질적 재정분권, 예산과정의 투명성을 주장하였다. 다음 <표 1>은 선행연구들이 채택한 지방정부 재정건전성의 영향요인들을 정리한 표이다.

선행연구들은 주로 재정자립도 및 자주도, 세입과 세출 등 재정적 요인이 지방재정건전성에 미치는 영향들을 살펴보았다. 반면, 각 지방정부들이 갖는 재정적 요인에 영향을 미칠 수 있는 근본적인 요인인 인구 통계학적, 경제·사회적 환경 등의 영향에 대해서는 큰 비중을 두어 연구하지 않았다. 다만, 윤석완(2010)의 연구는 지방세수와 지방정부의 인구특성을 연관하여 지역의 인구 및 노령인구와 지방재정수지의 관계를 분석하였다. 하지만, 해당 연구는 인구구조와 재정건전성 사이의 관계를 규명하였지만, 주요 변수와 재정수지 사이의 관계만을 실증분석 하였으며, 지방재정건전성에 영향을 미칠 수 있는 다른 영향요인들에 대한 분석이 부족하다는 한계를 갖는다. 본 연구는 이러한 선행연구들의 논의에 더하여 지방자치단체의 재정적 요인에 근본적으로 영향을 미칠 수 있는 요인을 확인하기 위해 사회·경제적 요인과 지방재정건전성 사이의 관계를 실증분석 하고자 한다.

〈표 1〉 우리나라 지방정부 재정건전성에 영향을 미치는 요인

| 저자 | 독립변수 | 종속변수 |
|--------------------|---|--------------------|
| 오영호, 편도창 (2012) | 재정자립도, 재정자주도, 1인당 자체수입, 경상수지비율, 세입세출총당비율, 지방채상환비율, 재정계획운영비율, 세입예산반영비율, 투자비율 | 1인당 부채규모, 부채비율 |
| 위정복 (2015) | 재정자립수준, 지방세수입, 투자지출, 복지지출 | 통합재정수지, 채무부담증가 |
| 이성근 외 (2016) | 세입: 재정자주도, 재정력지수 세출: 행사축제성경비비율, 민간단체보조금비율, 복지비비율, 인건비비율 | 예산대비 채무비율 |
| 최예나, 김상헌 (2017) | 재정자립도, 복지지출비중, 지역개발지출 비중, 노령인구 비중, 인구 1,000명당 종사자수 | 통합재정수지비율, 지자체부채비율 |
| 이현정, 조정래 (2019) | 지방세 비과세·감면 비율 | 재정수지, 예산대비 부채비율 |
| 강경표, 김준기 (2020) | 의회예산권한지수: 의회예산 수정정도, 예산안 심사기간, 예결위 개최횟수, 의회경비, 예비비, 전임의원 비율 | 통합재정수지비율, 지방채무 증가율 |

Ⅲ. 연구설계

1. 변수 설정

앞서 본 연구는 지방정부의 재정건전성을 기계학습의 목표변수(종속변수)로 하며, 목표변수에 영향을 주는 특성(독립변수)들은 인구 통계학적 변화와 사회·경제학적 요인을 수집하여 분석하기로 하였다. 지방정부의 재정건전성을 변수로 수집하기 위해 본 연구에서는 선행연구에서 지방재정건전성을 판단하였던 통합재정수지비율, 예산대비채무비율, 채무상환비율을 목표변수로 설정하며, 이외 특성변수는 지방재정건전성에 영향을 미칠 수 있는 인구사회학적 변수와 사회·경제환경 변수로 구성하는데 이때, 인구사회학적 변수는 지방소멸지수로, 사회·경제환경 변수는 UN의 기능별 분류에 따른 COFOG 정부기능분류 기준을 따라 변수를 구성한다.

1) 변수 설정 및 구성

(1) 목표변수(target, 종속변수)

본 연구의 목표변수는 지방정부의 재정건전성이다. 본 연구는 지방정부의 재정건전성을 측정하기 위해 지방재정 분석 지표를 활용하고자 한다. 2020년 행정안전부의 지방재정분석 지표에 따르면 재정건전성은 통합재정수지비율, 관리채무비율, 통합유동부채비율, 공기업부채비율로 나눌 수 있다. 여기서 통합재정수지비율과 관리채무비율은 통합회계를 대상으로 측정 및 관리되는 지표이며, 통합유동부채비율과 공기업부채비율은 공사, 공단, 출자출연 등의 회계를 통해 관리되는 지표이다.

본 연구는 데이터의 통일성을 위해 통합회계를 통해 관리되는 통합재정수지비율을 설정한다. 다만, 관리채무비율의 경우, 자료 확보가 어려워 지방재정건전성의 다른 지표로 활용되는 예산대비채무비율과 채무상환비율²⁾을 설정하였다. 목표변수는 재정정보공개시스템 열린재정(<https://www.openfiscaldata.go.kr/op/ko/index>)과 지방재정365(<https://lofin.mois.go.kr/portal/main.do>)를 통해 수집하였다.

(2) 특성변수(feature, 독립변수)

본 연구의 특성변수는 인구구조의 변화와 사회·경제적 환경이다. 인구구조의 변화는 지방소멸지수를 활용하여 측정한다. 지방소멸지수란 20~39세 여성 인구를 65세 이상 인구로 나눈 값으로 인구의 유출·유입 등 다른 변수가 크게 작용하지 않을 경우 약 30년 뒤에 해당 지역이 없어질 가능성이 높다는 의미를 가진다. 해당 지수는 노령층의 인구변화와 출생률 및 아동·청소년 등의 변화를 유추할 수 있기 때문에 해당 지표를 통해 인구구조의 변화를 측정하고자 한다.

2) 지방재정분석과 유사한 기능을 하는 지방재정 위기관리제도는 채무관리를 평가하기 위해 예산대비채무비율과 채무상환비율을 지표로 설정하고 있다.

다음으로, 지방정부의 사회·경제적 환경을 확인하기 위해 정부기능별분류(이하 COFOG; Classification of the Functions of Government)에 따라 사회·경제적 환경을 구분하여 분석하고자 한다. COFOG는 일반공공행정, 국방, 공공질서 및 안전, 경제활동, 환경보호, 주거 및 지역사회 시설, 보건, 휴양·문화·종교, 교육, 사회보호 총 10가지의 분야³⁾로 이루어져 있으며, 이를 통해 지방정부 전체적인 사회·경제적 환경을 확인할 수 있다. 분석에 필요한 변수들은 국가통계포털(KOSIS, <http://kosis.kr/>)을 통해 수집하였다. 다음 <표 2>는 본 연구의 변수 구성을 정리한 표이다.

<표 2> 변수의 구성⁴⁾

| 특성 | 자료구성 | | 특성 | 자료구성 |
|------------------------------------|---|--------------------------------|------------------|---|
| 재정건전성 지표 (목표값, target value) | y_1 | 통합재정수지비율 | 휴양 ·문화 ·종교 | 십만 명 당 문화기반시설 수, 문화기반시설 수 |
| | y_2 | 예산대비채무비율(BTL ⁵⁾ 포함) | | |
| | y_3 | 채무상환비율 | | |
| 지방소멸지수 | 20~39세 여성 인구 수/65세 이상 인구 수 | | 경제활동 | 각 사업의 사업체 수, ⁶⁾ 종사자 수 |
| 일반공공행정 | 혼인건수, 조혼인율, 합계출산율, 평균연령, 남녀성비, 인구증가율, 당해 년 총 인구, 전년총인구, 천 명당 외국인 수, 주민등록인구, 출생아수, 가구 수, 일반공공행정예산비중, 일반공공행정분야예산액, 토지거래면적 | | 사회보호 | 고위험음주율, 비만율, EQ.5D(건강상태 표준화), 주관적건강수준인지율, 인구 십만 명 당 자살률, 건강보험 적용인구 현황, 노인 천 명 당 노인여가복지시설 수, 노인여가 복지시설 수, 60세 이상 주민등록인구, 사회복지예산비중, 사회복지분야예산액 |
| 주거 및 지역사회건설 | 하수도보급률, 상수도보급률, 교통문화지수, 운전행태영역, 교통안전영역, 보행행태영역, 1인당 자동차등록대수, 도시지역면적, 주택 수 | | 환경보호 | 녹지지역면적, 녹지지역면적비율 |
| | | | 공공질서 및 안전 | 교통사고, 화재, 범죄, 자연재해, 생활안전, 자살, 감염병의 각 등급 |
| 보건 | 병원 수, 보건소 수, 상급종합병원 수, 약국 수, 요양병원 수, 의원 수, 종합병원 수, 치과병원 수, 한방병원 수, 천 명 당 의료기관병상수, 총 병상 수, 보건분야예산액 | | 교육 | 교원 1인 당 학생 수, 재직 학생 수, 교원 수, 유치원교원 수, 유치원 수, 유치원아 수, 천 명 당 사설학원 수, 사설학원 수, 초등교원 수, 초등학교수, 유치원학급 당 학생 수, 초등학교학급 당 학생 수, 중학교학급당 학생 수, 고등학교학급 당 학생 수 |

더하여 변수의 기술통계와 상관관계 분석결과는 부록에서 확인할 수 있으며, 변수의 대치(Imputation, Moving Average 대치방법을 사용) 및 전처리 과정에서 종속변수 등의 원천 자료의 부족재로 인해 제외된 데이터는 다음과 같다.

- 3) 국방 분야의 경우, 국가 단위에서만 확인 가능한 지표이므로 해당 변수는 분석에서 제외하였다.
- 4) 목표변수인 재정건전성 외 모두 특성변수의 구성이다.
- 5) 민간투자사업
- 6) 농업, 광업, 제조업, 건설업, 도소매업, 운수창고업, 숙박음식점업, 전기가스증기 및 공기조절공급업, 정보통신업, 금융보험업, 수도하수폐기물원료재생업, 부동산업, 전문과학기술서비스업 등의 사업체 수

〈표 3〉 변수 전처리 과정 설명과 전처리 후 제외된 시군구

| 자료 전처리 방법(Imputation) |
|---|
| 1) 시도 기준 수집으로 인해 시군구 기준 필터링 시 전체 결측 발생 변수 제거 2) 전체 수집 자료 중 분석에 사용할 2015~2021 자료만 추출 3) imputation: 파이썬의 imputation 라이브러리 중 impute 라이브러리 사용. impute 라이브러리의 함수 중 이동 평균방법을 사용하는 moving_window() 함수를 사용하여 각 시군구별 imputation을 진행 4) 결측 자료가 전체 자료의 5%이상 되는 변수 제거: 결측된 자료를 포함한 행을 제거하는 것 보다 열을 제거하는 것이 자료 손실이 적게 발생 5) 결측 자료를 포함한 행 제거 |
| 제외 시군구 |
| 서울 본청, 부산 본청, 부산 서구, 대구 본청, 대구 서구, 대구 달성군, 인천 본청, 인천 동구, 인천 옹진군, 광주 본청, 대전 본청, 대전 동구, 울산 본청, 세종 본청, 경기 본청, 경기 수원시, 경기 성남시, 경기 고양시, 경기 용인시, 경기 안산시, 경기 안양시, 경기 가평군, 강원 본청, 충북 본청, 충북 청주시, 충남 본청, 전북 본청, 전북 전주시, 전북 고창군, 전남 본청, 전남 신안군, 경북 본청, 경북 포항시, 경북 경산시, 경북 울릉군, 경남 본청, 경남 창원시, 제주 본청 |

2. 분석 방법

1) 변수 변환과 기계학습 모형 구축 방법

(1) 연속형 목표변수 범주화

본 연구의 연구목적은 지방정부의 재정건전성에 미치는 영향요인을 파악하기 위함으로 해당 영향요인의 중요도(importance)를 구하여 그 영향정도를 파악할 뿐 통계적·수학적으로 엄격한 가정을 통한 회귀계수(β)를 추정하지 않는다. 따라서 본 연구에서는 목표변수인 지방재정건전성의 3가지 변수를 각각 평균이상(1), 평균이하(0) 범주형 변수로 변환하여 분류모형(classification)을 통한 지방재정건전성에 영향을 미치는 요인을 추정하고자 한다. 이는 앞서 서술하였듯이 기계학습에서 회귀예측(regression)을 하는 방법은 본 연구에서 의도하는 영향요인 탐색과는 거리가 멀뿐만 아니라, 다수의 요인이 작용하는 기계학습모형으로 예측값을 추정하는 경우 예측값의 정확도가 높다고 하여도 어떠한 특성변수가 예측값에 작용하는지 정확히 추정할 수 없기 때문에 본 연구에서는 적합하지 않다. 따라서, 본 연구에서는 지방재정건전성의 값을 평균이상과 평균이하로 나누어 해당 범주형에 미치는 특성요인의 중요도(feature importance)만을 추정하여 보고한다.

(2) 활용 기계학습 모형과 방법

본 연구에서 활용하는 기계학습 모형은 로지스틱 회귀분석, SVM, DT(Decision Tree), RF(Random Forest), Boosting 등의 모형이 활용된다.⁷⁾ 해당 모형은 앞서 서술한 재정건전성 평균값을 기준으로 하여 범주형 변수로 전환한 목표변수를 활용하기 때문에 분류모형으로 목표변수에 미치는 특성들의 영향을 파악하기 위하여 구축한다. 각 모형에 대한 간단한 설명은 아래와 같다.⁸⁾

7) 다수의 모형을 구축하는 이유는 하이퍼파라미터(hyperparameter)의 세부조정을 통해 모형의 정확도를 높일 경우

8) 본 연구에서는 연구의 목적에 따라 모형에 적용하는 상세한 함수, 의사코드 등에 대한 설명은 생략하기로

① 로지스틱 회귀분석

로지스틱 회귀분석은 종속변수가 0 또는 1로 이분화(二分化) 되어있을 때 사용할 수 있는 일종의 회귀분석이다. 로지스틱 회귀분석은 종속변수에 미치는 독립변수의 영향을 계수를 통해 추정할 수 있으며, 독립변수의 선형 결합을 이용하여 목표변수가 어떠한 범주에 속할지를 예측하는데 사용될 수 있다. 이때 목표변수의 범주 예측값은 0과 1사이로 예측되며 시그모이드 함수를 적용하여 0.5값을 기준으로 0과 1로 나누게 되어 각 집단에 속하게 된다. 이러한 방식으로 본 연구에서는 지방재정건전성의 평균이상(1)과 평균이하(0)를 예측한다.

② Support Vector Machine(SVM)

서포트벡터머신은 자료를 분류(classification)하는 알고리즘으로 다차원 공간에서 초평면을 구성하며, 이차원 공간에서는 이진 분류 자료 혹은 그 이상의 자료를 구분 지어 나누는데 구분선(선형)이고, 비선형은 3차원 이상의 고차원적 구분면을 도출한다(Suthaharan, 2016). 본 연구에서는 이러한 서포트벡터머신을 활용하여 목표변수를 분류하는데 활용한다.

③ 의사결정나무(Decision Tree)

의사결정나무는 CART, CHAID 등과 같은 알고리즘을 활용하여 목표변수가 특성변수의 값에 따라 나뉘는 구조이며, 뿌리-중간-끝 마디의 나무구조로 표현할 수 있다(Myles *et al.*, 2004). 이러한 의사결정나무는 시각화를 통해 분류 기준을 확인할 수 있으며, 분류에 대한 설명 기준을 쉽게 확인할 수 있다. 본 연구에서는 분류나무에서 지니지수를 활용하는 CART 알고리즘을 사용하여 모형을 구축하고 예측을 진행한다.

④ 랜덤포레스트(Random Forest)

랜덤포레스트 모형은 배깅의 일종으로 복원추출하여 동일한 크기의 자료 집합을 다수 구성하여 여러 모형을 학습하는 모형으로 나무 분류 방법(Tree-type classification)의 앙상블 모형이다. 랜덤포레스트는 의사결정나무(CART-like tree)를 훈련과정에서 다수 학습모형으로 구축하고, 각 모형마다 자료의 부트스트랩 샘플(bootstrapped sample)을 구성하고 이렇게 구성된 무작위적 자료에서 입력 변수가 적용되어 분류 모형을 구축한다(Breiman *et al.*, 1984; Breiman, 2001).

⑤ GBM(Gradient Boosting Machine)

부스팅 모형은 앞서 서술한 랜덤포레스트에서 설명한 배깅과 비슷한 방법인 앙상블(ensemble) 기반 모형으로 다수의 분류기를 생성하여 그 예측 성능을 높여가는 방법이다. 배깅과의 차이점은 다수의 분류기가 순차적 생성이 되기 때문에 이전 분류기의 학습이 다음 분류기 학습에 영향을 미치는 점이다. 학습 과정에서 다차원 공간의 분류경계선(borderline)의 데이터 가중치가 증가하여 과적합에는 강하지만 분류결과에 대한 해석이 불가능하다는 단점이 있으나 예측의 성능을 높이기

하며, 각 모형에 대한 이해를 할 수 있는 수준에서 서술하기로 한다.

위한 방법으로 본 연구에서도 활용한다. 본 연구에서는 이러한 부스팅 모형 중 Gradient Boosting 모형을 구축하여 분석한다.

본 연구에서는 앞서 서술한 바와 같이 LR(Logistic Regression), SVM(Support Vector Machine), DT(Decision Tree), RF(Random Forest), Boosting 모형을 구축하여 목표변수에 영향을 주는 특성변수를 도출하며, 머신러닝 분석방법에서 통상적으로 사용하는 홀드아웃(hold-out) 방법을 통해 자료를 훈련자료(train set)와 시험자료(test set)로 나누어 분석한다. 이때 목표변수는 회귀분석에서 종속변수로 볼 수 있으며, 특성변수는 독립변수와 같다. 홀드아웃 방법은 머신러닝에서 훈련자료와 시험자료를 분리하는 방법의 일종으로 원 자료로부터 훈련자료와 시험자료를 7:3, 8:2 등의 비율로 무작위 분리하여 생성하는 교차 검증방법이다. 훈련자료는 머신러닝 모델링에 이용되며, 시험자료는 머신러닝을 평가하기 위해 사용되는 자료이다. 본 연구에서는 훈련자료와 시험자료의 비율은 7:3으로 구성하며, 훈련자료로 모형을 구축하고, 시험자료를 통한 모형의 성능을 측정한다. 이때 혼동행렬(Confusion Matrix)⁹⁾을 이용한 정확도(accuracy)¹⁰⁾를 확인한다. 이러한 분석은 파이썬(구글 코랩), R(R studio) 등의 소프트웨어를 사용하여 분석하였다.

IV. 분석 결과

1. 머신러닝 분석결과

1) 목표변수 1: 통합재정수지비율

〈표 4〉 통합재정수지비율이 목표변수인 각 모델의 성능 결과

| 구분 | LR | SVM | DT | RF | GBM |
|------------------|--|---|---|---|---|
| Accuracy | 0.5977 | 0.593 | 0.593 | 0.6256 | 0.5767 |
| Error rate | 0.4023 | 0.407 | 0.407 | 0.3744 | 0.4233 |
| Sensitivity | 0.6357 | 0.6186 | 0.6029 | 0.6471 | 0.6000 |
| Specificity | 0.5407 | 0.5619 | 0.5759 | 0.5990 | 0.5459 |
| confusion matrix | Reference Prediction 0 1 0 164 79 1 94 93 | Reference Prediction 0 1 0 146 85 1 90 109 | Reference Prediction 0 1 0 164 67 1 108 91 | Reference Prediction 0 1 0 154 77 1 84 115 | Reference Prediction 0 1 0 147 84 1 98 101 |

혼동행렬표(confusion matrix)

| | | | |
|----|----------|--------------------|--------------------|
| | | 예측 | |
| | | 예측 분류 \hat{y} | |
| 실제 | Negative | True Negative(TN) | False Positive(FP) |
| | Positive | False Negative(FN) | True Positive(TP) |

9)

$$10) accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

통합재정수지비율(y_1)을 목표변수로 분석한 각 모형의 성능 중 랜덤포레스트 모형이 가장 성능이 좋았으며 62.56%의 정확도를 보였다. 랜덤포레스트 모형에서 도출된 통합재정수지비율에 영향을 미치는 상위 20개 요인은 아래 <표 5>와 같다.

<표 5> RF 모형 특성중요도(종속변수: 통합재정수지비율)

| 상위 20개 특성중요도 변수 | feature importance |
|----------------------------------|--------------------|
| COFOG_일반공공행정8_인구증가율 | 100 |
| COFOG_일반공공행정11_일반공공행정예산비중 | 68.96 |
| COFOG_주거및지역사회건설8_교통안전영역 | 66.87 |
| COFOG_사회보호9_사회복지예산비중 | 40.89 |
| COFOG_주거및지역사회건설7_운전행태영역 | 37.35 |
| COFOG_경제활동10_농업임업및어업_종사자수 | 31.65 |
| COFOG_보건8_치과병원 | 26.81 |
| COFOG_교육11_유치원학급당_학생수 | 21.55 |
| COFOG_공공질서및안전3_범죄 | 19.93 |
| COFOG_경제활동28_전기가스증기및공기조절공급업_종사자수 | 19.88 |
| COFOG_공공질서및안전6_자살 | 18.39 |
| COFOG_경제활동24_예술스포츠및여가관련서비스업_종사자수 | 17.15 |
| COFOG_주거및지역사회건설6_교통문화지수 | 16.78 |
| COFOG_경제활동25_운수및창고업_사업체수 | 16.13 |
| COFOG_일반공공행정6_평균연령 | 15.19 |
| COFOG_경제활동4_광업_종사자수 | 15.14 |
| COFOG_교육14_고등학교학급당_학생수 | 15.09 |
| COFOG_경제활동1_건설업_사업체수 | 14.99 |
| COFOG_보건7_종합병원 | 14.92 |
| COFOG_일반공공행정4_조혼인율 | 12.21 |

2) 목표변수 2: 예산대비채무비율(BTL 포함)

<표 6> 예산대비채무비율이 목표변수인 각 모형의 성능 결과

| 구분 | LR | SVM | DT | RF | GBM |
|------------------|--|--|--|--|--|
| Accuracy | 0.7395 | 0.7814 | 0.7256 | 0.807 | 0.814 |
| Error rate | 0.4023 | 0.2186 | 0.2744 | 0.193 | 0.186 |
| Sensitivity | 0.7929 | 0.8152 | 0.8175 | 0.8117 | 0.8339 |
| Specificity | 0.6033 | 0.7008 | 0.5808 | 0.7925 | 0.7642 |
| confusion matrix | Reference Prediction 0 1 0 245 48 1 64 73 | Reference Prediction 0 1 0 247 38 1 56 89 | Reference Prediction 0 1 0 215 70 1 48 97 | Reference Prediction 0 1 0 263 22 1 61 84 | Reference Prediction 0 1 0 256 29 1 51 94 |

예산대비채무비율(y_2)을 목표변수로 분석한 각 모형의 성능 중 부스팅 GBM 모형이 가장 성능이 좋았으며 81.4%의 정확도를 보였다. 모형에서 도출된 통합재정수지비율에 영향을 미치는 상위 20개 요인은 아래 <표 7>과 같다.

<표 7> GBM 모형 특성중요도(종속변수: 예산대비채무비율)

| 상위 20개 특성중요도 변수 | feature importance |
|-----------------------------------|--------------------|
| COFOG_일반공공행정8_인구증가율 | 100 |
| COFOG_경제활동4_광업_종사자수 | 88.23 |
| COFOG_주거및지역사회건설5_상수도보급률 | 82.16 |
| COFOG_사회보호6_노인천명당_노인여가복지시설수 | 74.59 |
| COFOG_주거및지역사회건설7_운전행태영역 | 66.51 |
| COFOG_교육11_유치원학급당_학생수 | 65.85 |
| COFOG_주거및지역사회건설9_보행행태영역 | 64.63 |
| COFOG_사회보호9_사회복지예산비중 | 62.66 |
| COFOG_일반공공행정3_합계출산율 | 62.13 |
| COFOG_교육13_중학교학급당_학생수 | 51.86 |
| COFOG_사회보호2_고위험음주율 | 49.87 |
| COFOG_환경보호2_녹지지역면적비율 | 49.37 |
| COFOG_경제활동38_출판영상방송통신및정보서비스업_종사자수 | 45.64 |
| COFOG_경제활동44_공공행정국방및사회보장행정_종사자수 | 44.18 |
| COFOG_일반공공행정12_일반공공행정분야예산액 | 42.98 |
| COFOG_환경보호1_녹지지역면적 | 38.19 |
| COFOG_사회보호3_비만율 | 35.28 |
| COFOG_일반공공행정11_일반공공행정예산비중 | 33.87 |
| COFOG_일반공공행정7_남녀성비 | 32.58 |
| COFOG_공공질서및안전3_범죄 | 30.91 |

3) 목표변수 3: 채무상환비비율

<표 8> 채무상환비비율이 목표변수인 각 모형의 성능 결과

| 구분 | LR | SVM | DT | RF | GBM | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|------------------|---|--------|-----------|--------|------------|---|---|---|-----|----|---|----|----|---|--|-----------|--|------------|---|---|---|-----|----|---|----|----|---|--|-----------|--|------------|---|---|---|-----|----|---|----|----|---|--|-----------|--|------------|---|---|---|-----|----|---|----|----|--|--|-----------|--|------------|---|---|---|-----|----|---|----|-----|
| Accuracy | 0.7698 | 0.7837 | 0.7558 | 0.814 | 0.8326 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Error rate | 0.4023 | 0.2163 | 0.2442 | 0.186 | 0.1674 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Sensitivity | 0.7829 | 0.8113 | 0.7803 | 0.8211 | 0.8404 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Specificity | 0.7381 | 0.7188 | 0.6897 | 0.7949 | 0.8130 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| confusion matrix | <table border="1"> <thead> <tr> <th></th> <th colspan="2">Reference</th> </tr> <tr> <th>Prediction</th> <th>0</th> <th>1</th> </tr> </thead> <tbody> <tr> <th>0</th> <td>238</td> <td>33</td> </tr> <tr> <th>1</th> <td>66</td> <td>93</td> </tr> </tbody> </table> | | Reference | | Prediction | 0 | 1 | 0 | 238 | 33 | 1 | 66 | 93 | <table border="1"> <thead> <tr> <th></th> <th colspan="2">Reference</th> </tr> <tr> <th>Prediction</th> <th>0</th> <th>1</th> </tr> </thead> <tbody> <tr> <th>0</th> <td>245</td> <td>36</td> </tr> <tr> <th>1</th> <td>57</td> <td>92</td> </tr> </tbody> </table> | | Reference | | Prediction | 0 | 1 | 0 | 245 | 36 | 1 | 57 | 92 | <table border="1"> <thead> <tr> <th></th> <th colspan="2">Reference</th> </tr> <tr> <th>Prediction</th> <th>0</th> <th>1</th> </tr> </thead> <tbody> <tr> <th>0</th> <td>245</td> <td>36</td> </tr> <tr> <th>1</th> <td>69</td> <td>80</td> </tr> </tbody> </table> | | Reference | | Prediction | 0 | 1 | 0 | 245 | 36 | 1 | 69 | 80 | <table border="1"> <thead> <tr> <th></th> <th colspan="2">Reference</th> </tr> <tr> <th>Prediction</th> <th>0</th> <th>1</th> </tr> </thead> <tbody> <tr> <th>0</th> <td>257</td> <td>24</td> </tr> <tr> <th>1</th> <td>56</td> <td>93</td> </tr> </tbody> </table> | | Reference | | Prediction | 0 | 1 | 0 | 257 | 24 | 1 | 56 | 93 | <table border="1"> <thead> <tr> <th></th> <th colspan="2">Reference</th> </tr> <tr> <th>Prediction</th> <th>0</th> <th>1</th> </tr> </thead> <tbody> <tr> <th>0</th> <td>258</td> <td>23</td> </tr> <tr> <th>1</th> <td>49</td> <td>100</td> </tr> </tbody> </table> | | Reference | | Prediction | 0 | 1 | 0 | 258 | 23 | 1 | 49 | 100 |
| | Reference | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Prediction | 0 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 238 | 33 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 66 | 93 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | Reference | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Prediction | 0 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 245 | 36 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 57 | 92 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | Reference | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Prediction | 0 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 245 | 36 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 69 | 80 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | Reference | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Prediction | 0 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 257 | 24 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 56 | 93 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | Reference | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Prediction | 0 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 258 | 23 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 49 | 100 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

채무상환비율(y_3)을 목표변수로 분석한 각 모형의 성능은 앞서 예산대비채무비율에서 좋은 성능을 보여주었던 부스팅 GBM 모형이 마찬가지로 가장 성능이 좋았으며 83.26%의 정확도를 보였다. 모형에서 도출된 통합재정수지비율에 영향을 미치는 상위 20개 요인은 아래 <표 9>와 같다.

<표 9> GBM 모형 특성중요도(종속변수: 채무상환비율)

| 상위 20개 특성중요도 변수 | feature importance |
|-------------------------------------|--------------------|
| COFOG_사회보호6_노인천명당_노인여가복지시설수 | 100 |
| COFOG_경제활동27_전기가스증기및공기조절공급업_사업체수 | 73.59 |
| COFOG_환경보호2_녹지지역면적비율 | 65.35 |
| COFOG_일반공공행정7_남녀성비 | 61.09 |
| COFOG_사회보호9_사회복지예산비중 | 48.35 |
| COFOG_주거및지역사회건설5_상수도보급률 | 47.16 |
| COFOG_일반공공행정3_합계출산율 | 47.12 |
| COFOG_교육13_중학교학급당_학생수 | 40.29 |
| COFOG_사회보호2_고위험층주울 | 40.08 |
| COFOG_경제활동39_하수폐기물처리원료재생및환경복원업_사업체수 | 39.45 |
| COFOG_일반공공행정12_일반공공행정분야예산액 | 38.63 |
| COFOG_주거및지역사회건설9_보행행태영역 | 27.8 |
| COFOG_경제활동4_광업_종사자수 | 25.29 |
| COFOG_일반공공행정13_천명당_외국인수 | 25.26 |
| COFOG_보건10_보건분야예산액 | 25.16 |
| COFOG_경제활동44_공공행정국방및사회보장행정_종사자수 | 24.86 |
| COFOG_주거및지역사회건설10_인당_자동차등록대수 | 24.41 |
| COFOG_경제활동38_출판영상방송통신및정보서비스업_종사자수 | 24.16 |
| COFOG_환경보호1_녹지지역면적 | 21.61 |
| COFOG_교육14_고등학교학급당_학생수 | 20.59 |

2. 각 종속변수에 미치는 영향요인 정리와 종합결과

<표 10>의 분석결과에 따라 머신러닝 모형으로 도출한 특성중요도 상위 20개 변수와 함께 로지스틱회귀분석 결과를 함께 도출함으로써 특성의 중요도와 계수의 방향(+, -)을 확인할 수 있었다. 이러한 결과에 따라 경제활동(산업구조)과 일반공공행정(예산비중, 인구 구조 관련) 변수들이 다수가 지방재정에 영향을 미치고 있음을 확인하여 선행연구와의 일치를 확인할 수 있었다. 사회보호, SOC, 교육 등도 중요한 요인으로 확인할 수 있다.

〈표 10〉 각 종속변수에 영향을 미치는 특성중요도 상위 20개 변수 종합 결과 제시

| 통합재정수지비율 | 예산대비채무비율 | 채무상환비율 |
|--|--|---|
| 경제활동1_건설업_사업체수 (-0.058) | 경제활동38_출판영상방송통신및정보 서비스업_종사자수 (0.227) | 경제활동27_전기가스증기및공기조절 공급업_사업체수 (-0.019) |
| 경제활동10_농업임업및어업_종사자수 (0.004) | 경제활동4_광업_종사자수 (0.016) | 경제활동38_출판영상방송통신및정보 서비스업_종사자수 (0.177) |
| 경제활동24_예술스포츠및여가관련서 비스업_종사자수 (-0.074) | 경제활동44_공공행정국방및사회보장 행정_종사자수 (-0.139)** | 경제활동39_하수폐기물처리원료재생 및환경복원업_사업체수 (0.234)** |
| 경제활동25_운수및창고업_사업체수 (-0.055) | 공공질서및안전3_범죄 (-0.007) | 경제활동4_광업_종사자수 (0.006) |
| 경제활동28_전기가스증기및공기조절 공급업_종사자수 (-0.009) | 교육11_유치원학급당_학생수 (0.077)* | 경제활동44_공공행정국방및사회보장 행정_종사자수 (-0.161)*** |
| 경제활동4_광업_종사자수 (0.006) | 교육13_중학교학급당_학생수 (0.084) | 교육13_중학교학급당_학생수 (-0.026) |
| 공공질서및안전3_범죄 (0.025) | 사회보호2_고위험음주율 (0.001) | 교육14_고등학교학급당_학생수 (0.022) |
| 공공질서및안전6_자살 (0.033) | 사회보호3_비만율 (-0.017) | 보건10_보건분야예산액 (-0.056) |
| 교육11_유치원학급당_학생수 (0.022) | 사회보호6_노인천명당_노인여가복지 시설수 (0.039) | 사회보호2_고위험음주율 (0.007) |
| 교육14_고등학교학급당_학생수 (0.022) | 사회보호9_사회복지예산비중 (-0.267)*** | 사회보호6_노인천명당_노인여가복지 시설수 (-0.001) |
| 보건7_종합병원 (-0.065)* | 일반공공행정11_일반공공행정예산비중 (-0.019) | 사회보호9_사회복지예산비중 (-0.106)* |
| 보건8_치과병원 (0.082)* | 일반공공행정12_일반공공행정분야예 산액 (0.018) | 일반공공행정12_일반공공행정분야예 산액 (-0.007) |
| 사회보호9_사회복지예산비중 (-0.106)* | 일반공공행정3_합계출산율 (0.012) | 일반공공행정13_천명당_외국인수 (-0.036) |
| 일반공공행정11_일반공공행정예산비중 (0.015) | 일반공공행정7_남녀성비 (0.055) | 일반공공행정3_합계출산율 (-0.034) |
| 일반공공행정4_조혼인율 (-0.018) | 일반공공행정8_인구증가율 (0.004) | 일반공공행정7_남녀성비 (0.001) |
| 일반공공행정6_평균연령 (-0.206)* | 주거및지역사회건설5_상수도보급률 (-0.013) | 주거및지역사회건설10_인당_자동차등 등록대수 (-0.028) |
| 일반공공행정8_인구증가율 (0.018) | 주거및지역사회건설7_운전행태영역 (-0.039) | 주거및지역사회건설5_상수도보급률 (0.083)*** |
| 주거및지역사회건설6_교통문화지수 (-0.027) | 주거및지역사회건설9_보행행태영역 (-0.067) | 주거및지역사회건설9_보행행태영역 (-0.043) |
| 주거및지역사회건설7_운전행태영역 (0.015) | 환경보호1_녹지지역면적 (-0.050)* | 환경보호1_녹지지역면적 (-0.008) |

| | | |
|---|--|---|
| 주거및지역사회건설8_교통안전영역 (0.047) | 환경보호2_녹지지역면적비율 (0.030) | 환경보호2_녹지지역면적비율 (0.024) |
| 영향 요인 종합 | | |
| 경제활동(7개) 공공질서 및 안전(2개) 교육(2개) 보건(2개) 사회보호(1개) 일반공공행정(3개) 주거 및 지역사회 건설(3개) | 경제활동(3개) 공공질서 및 안전(1개) 교육(2개) 사회보호(4개) 일반공공행정(5개) 주거및지역사회건설(3개) 환경보호(2개) | 경제활동(5개) 교육(2개) 보건(1개) 사회보호(3개) 일반공공행정(4개) 주거및지역사회건설(2개) 환경보호(2개) |
| 경제활동(15개), 일반공공행정(12개), 사회보호(8개), 주거 및 지역사회 건설(8개), 교육(6개), 환경보호(4개) 공공질서 및 안전(3개), 보건(3개) | | |

주 1) 표 내용 내 괄호는 로지스틱 회귀분석에서 도출된 회귀계수(β)이며, 소수 넷째자리에서 반올림 처리하였음

2) 로지스틱 기준 유의수준: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

지방의 재정건전성에 영향을 미치는 요인의 해석은 다음과 같이 진행할 수 있다. 상기한 <표 10>의 내용의 특성들은 모형에서 목표변수의 범주를 분류하는데 중요하게 작용한 변수들이므로 모든 변수가 지방재정건전성에 영향을 미친다고 볼 수 있다. 추가적으로 각 목표변수에 따라 특성 중요도와 로지스틱 회귀분석 결과를 고려하여 제시하면, 통합재정수지비율의 경우 종합병원 수, 사회복지예산비중, 평균연령 등이 일부 음의 영향을 줄 수 있음으로 파악되었고, 예산대비채무비율에는 공공행정 종사자 수, 사회복지예산비중, 녹지지역면적 등이, 그리고 채무상환비율에는 공공행정 종사자 수, 사회복지예산비중 등이 영향을 통계적으로도 확인할 수 있다.

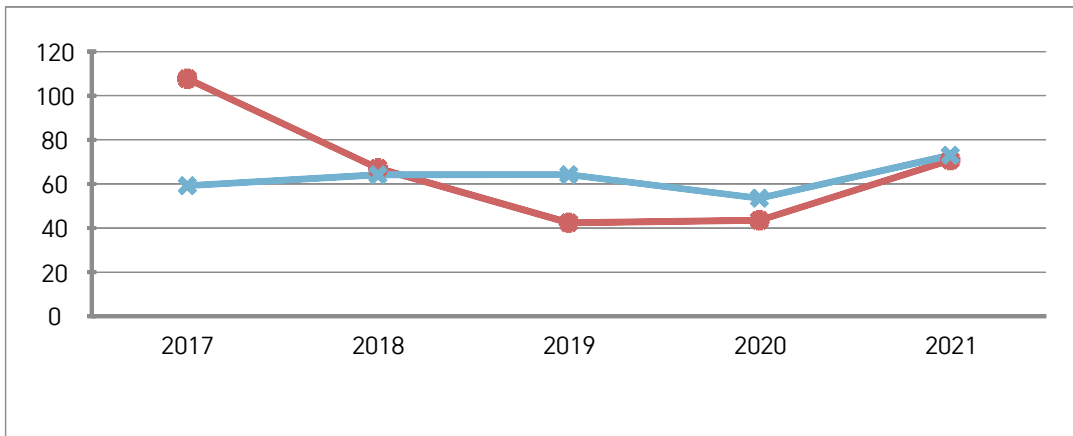
위의 결과를 종합해보면 근본적으로 재정건전성에 주요하게 영향을 주는 요인은 결국 경제적인 측면과 사회복지예산과 같은 사회복지 분야 요인으로 보인다. 먼저, 경제적 측면의 경우, 비록 통계적으로 유의미하게 나타나지는 않았지만, 지방자치단체의 통합재정수지비율에 중요한 영향을 미치는 변수로 파악되었다. 구체적으로, 농업 및 어업 종사자 수나 광업 종사자 수 등 1차 산업 분야는 통합재정수지에 양의 영향을 미치는 반면, 건설 사업체 수, 예술·스포츠 관련 서비스 종사자 수 등 2·3차 산업 분야의 경우, 음의 영향을 미치는 것으로 확인되었다. 농업 및 어업 등의 1차 산업 분야의 경우, 특별회계와 기금, 의무지출 사업 등 해당 목적에만 활용할 수 있는 예산이 마련되기 때문에 지방자치단체의 전체 수입 규모를 늘려 통합재정수지비율에 긍정적 영향을 미칠 수 있다. 반면, 대부분의 2·3차 산업 분야의 경우, 해당 분야를 위한 재원이 따로 마련되어있지 않으며, 사업 수행을 위해 상위정부로부터 보조금을 지원받더라도 지방자치단체의 지출 부담이 있기 때문에 통합재정수지에 음의 영향을 미칠 수 있다.

다음으로, 사회복지 예산의 비중이 증가하면 예산대비채무비율과 채무상환비율이 낮아지는 결과가 나타났다. 지방자치단체의 사회복지예산의 경우, 전체 예산에 평균 약 34%를¹¹⁾ 차지할 정도로 많은 비중을 차지하고 있다. 이에 더하여 지방자치단체의 사회복지 예산사업은 중앙정부에서 국고보조금을 지급하여 진행되는 사업이 많기 때문에 지방자치단체의 사회복지예산의 규모에

11) 2015년부터 2021년까지 기초지방자치단체의 전체 예산 대비 사회복지 분야 예산 비중의 평균값

큰 영향을 미칠 수 있고, 이는 곧 지방자치단체의 전체 예산 규모에도 영향을 미친다. 실제로, 지난 2017년부터 2021년까지 지방자치단체의 사회복지 분야 예산 중 국고보조금이 차지하는 비율은 평균 약 63%로 나타났으며, 다음 <그림 1>을 통해 사회복지분야 국고보조금 비율과 예산대비채무비율 사이에 반비례 관계를 확인할 수 있다. 이러한 결과는 지방자치단체의 예산대비채무비율에 영향을 미치는 요인들은 다양하지만, 사회복지분야예산의 비율이 전체 예산 규모를 증가시켜 예산대비채무비율에 음의 영향을 미칠 수 있다는 것을 보여준다.

<그림 1> 사회복지분야 국고보조금 비율, 예산대비채무비율 추이(2017~2021년)



자료: e나라도움(<https://www.gosims.go.kr/hg/hg001/retrieveMain.do>)과 열린재정(<https://www.openfiscaldata.go.kr/op/ko/index>)의 자료를 활용하여 저자가 작성.

동시에 채무상환비율의 경우, 앞서 서술하였던 중앙정부와 지방자치단체의 사회복지 예산의 지원에도 불구하고 지방자치단체의 예산이 일정비율로 투입되므로, 지방자치단체 자체의 재정적 역량이 부족한 경우 채무상환비율에 부정적인 영향을 미칠 수 있다. 이는 평균연령의 증가와 관련하여서 함께 고려되어야 할 부분이며, 지방소멸을 고려하여 지방재정건전성을 확보해야한다. 지방소멸이 실제로 영향을 미치고 있는지에 대해 모형의 정확도와는 별개로 각 모형에서 지방소멸지수의 중요도 파악과 재정건전성에 긍정·부정적 영향을 미칠 수 있는지 다음 목에서 확인한다.

3. 지방소멸의 영향

본 연구에서는 지방소멸이 재정건전성에 미치는 영향이 있는지 파악하기 위해 인구사회학적 COFOG 지표 변수와 더불어 지방소멸지수를 특성변수로 구성하고 모형에 반영하였다. 각 모형을 도출한 후 모형에서 활용된 특성들의 중요도를 도출하여 지방소멸지수의 영향을 파악하였으며, 해당 결과는 아래 <표 11>과 같다.

〈표 11〉 지방소멸이 재정건전성에 미치는 영향 분석 결과

| 지방소멸지수 → 종속변수에 영향 | | | | | |
|-------------------|--------------|----------|------------|--------|--------------|
| 통합재정수지비율 | | 예산대비채무비율 | | 채무상환비율 | |
| LR | 계수: -0.171** | LR | 계수: -0.118 | LR | 계수: -0.171** |
| SVM | - | SVM | - | SVM | - |
| DT | - | DT | - | DT | 변수중요도: 34.49 |
| RF | - | RF | - | RF | - |
| GBM | 변수중요도: 14.20 | GBM | - | GBM | - |

주: $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

지방소멸지수가 각 종속변수에 주는 영향은 모든 종속변수에 대해서 음의 영향이 작용한다고 확인되었으나, 예산대비채무비율에서는 유의수준에 미치지 못하였기 때문에 통합재정수지비율과 채무상환비율에서만 영향을 파악할 수 있다. 지방소멸지수가 증가하는 경우는 지방소멸 가능성에서 멀어지는 경우(지방소멸에서 안전함=지방이 소멸되지 않음)인데, 계수가 모두 부적 영향으로 결과가 도출되었다.

이 결과는 지방소멸지수가 커지는 경우 오히려 재정건전성을 악화시킬 수 있음을 제시하고 있는데, 지방소멸지수가 커질 수 있는 가능성이 있는 경우는 가임기 여성인구의 증가, 노인인구의 감소이기 때문에 지방소멸과 재정건전성에서 자료에 영향을 주는 왜곡이 있을 수 있다고 파악된다. 이는 현재 지방자치단체에 대해 중앙이 지방에 주는 이전 재원의 역할이 크다고 파악된다. 결국 이는 지방자치단체가 지방소멸을 이유로 사회복지와 같이 이전재원을 많이 받아 예산과 정부 수입이 늘어남으로써 도출된 결과로 보인다.

본 연구의 분석결과, 경제활동(15개), 일반공공행정(12개), 사회보호(8개), 주거 및 지역사회 건설(8개), 교육(6개), 환경보호(4개) 공공질서 및 안전(3개), 보건(3개) 분야에서 다양한 요인이 재정건전성에 영향을 미칠 수 있으며, 지방소멸 또한 영향이 확인되었으나, 지방소멸의 경우 이미 지방소멸대응기금 등으로 재원이 배분되면서 재정건전성에는 오히려 긍정적인 영향을 미치고 있음으로 확인되었다고 볼 수 있다.

V. 결론 및 시사점

본 연구는 기존 선행연구들에서 확인하지 못했던 지방재정건전성과 인구통계학적, 사회·경제적 환경의 관계를 분석하였다. 분석결과, 통합재정수지비율은 종합병원 수, 사회복지예산비중, 평균연령에 음의 영향을 받았으며, 공공행정 종사자 수, 사회복지예산비중, 녹지지역면적 등은 예산대비채무비율에 음의 영향을 미쳤으며, 채무상환비율의 경우, 공공행정 종사자 수, 사회복지예산비중이 음의 영향을 미친 것을 통계적으로 확인하였다.

이러한 결과를 종합해보면 지방자치단체는 주로 사회복지와 관련한 영향을 가장 크게 받는 것

을 알 수 있다. 특히, 사회복지비의 경우, 채무상환비비율과 예산대비채무비율에 모두 음의 영향을 미치는 것으로 나타났으며, 이를 통해 예산 증가와 사회복지 예산의 비중은 반비례 관계를 가질 수 있다는 점이 고려되어야 한다. 동시에 채무상환비비율의 경우, 중앙정부와 지방자치단체의 사회복지 예산의 지원에도 불구하고 지방자치단체의 예산이 일정비율로 투입되므로 지방자치단체의 재정적 역량이 부족한 경우 채무상환비비율에 부정적인 영향을 미칠 수 있다고 해석할 수 있다.

경제적 측면의 경우, 지방자치단체의 통합재정수지비율에 영향을 미치는 것으로 나타났으며, 이를 로지스틱 회귀분석한 결과 통계적으로 유의미하지는 않으나, 농업 및 광업 등 1차 산업 분야 산업은 양의 영향을 미친 반면 문화 및 관광 또는 건설업 등의 2·3차 산업 분야는 음의 영향을 미친 것으로 나타났다. 1차 산업 분야와 달리 2·3차 산업 분야 산업의 경우, 특별회계나 기금, 의무지출 사업 등과 같이 해당 분야만을 지원할 수 재정적 통로가 없기 때문에 산업을 지원하기 위해서는 지방자치단체의 자체 예산이나, 상위정부로부터 이전재원을 활용해야 하며, 이는 곧 통합재정수지 악화에 영향을 줄 수 있다. 따라서 향후 지방자치단체는 장기적 관점에서 해당 분야의 사업체들이 자력으로 경제적 기반 산업을 유지할 수 있는 원동력을 확보할 수 있도록 지원해야 하며, 이를 위해서는 산업단지 유치, 세제 혜택 등을 통한 기업 유치 및 중소·중견기업 대상 지원을 통한 기업 이전 등 다양한 기업 및 산업 유치 노력이 필요하다고 보인다.

본 연구의 분석결과는 지방정부의 중앙정부의 이전재원에 대한 의존을 극명하게 보여준다고 할 수 있다. 즉, 중앙정부의 이전재원에 의지할수록 지방재정건전성이 향상되는 아이러니한 결과가 나타났다. 이러한 상황이 지속되면, 지방자치단체는 중앙정부가 유도하는 정책 방향으로 예산을 지출할 수밖에 없으며, 재정건전성을 위해 지역의 특성에 기반한 경제 환경조성을 위한 재원을 투자할 수 없는 상황에 직면할 수 있다. 특히, 우리나라처럼 지방자치단체의 보조사업 비중이 57%¹²⁾에 달하는 상황에서는 지방자치단체가 중앙정부로부터 재정적 자립을 할 수 있는 방안 마련이 필요하다. 이를 위해서 지방자치단체의 지방세율의 탄력적으로 운영하여 세금혜택 및 징세조정권한 등 지방세 조정권한을 확대하는 방법이 있을 수 있다.

하지만, 지방세 운영이 비탄력적으로 운영되는 현재 상황에서는 산업단지 유치, 세제 혜택 등을 통한 기업 유치 및 중소·중견기업 대상 지원을 통한 기업 이전 등 다양한 기업 및 산업 유치 노력 등 지방정부의 세수 확대를 위한 지방자치단체의 경제적 기반 활성화를 위해 노력해야 한다. 더하여 현재 지방소멸 문제가 극심해지고 있으므로 이를 고려하여 지방 이전 법인에 대한 파격적 지원을 해줌으로써 법인의 지방 이전 활성화가 필요하다. 현재 조세특례제한법 제63조의2를 통해 지방 이전 법인에 대한 세액을 감면해주고 있는 바, 해당 법령을 개정하여 세액 감면기간과 감면세액 혜택을 더욱 늘려 지방자치단체의 경제적 기반을 유지하는 방법을 강구해야 한다.

본 연구는 자료 전처리를 진행하면서 일부 지역이 배제된 점과 일부 자료가 완전하지 않아 대치법(imputation)을 활용하였기 때문에 자료가 컴퓨터의 계산에 의해 임의적으로 채워지면서 발생할 수 있는 왜곡에 대한 문제가 있을 수 있다는 점을 한계로 들 수 있다. 더하여 머신러닝 모형 중

12) 2021년 기준 지방자치단체의 보조사업 비중은 평균 57.3%이다.

5가지 모형만을 활용하였는데 이는 신경망모형, KNN 등이 사전 분석에서 성능이 높지 않음을 확인하고 선정하였기 때문에 연구자의 활용 모형 선택에 주관성이 있음을 제시할 수 있다.

이러한 한계에도 불구하고 본 연구에서는 앞선 선행연구에서 반영하지 않은 다양한 변수들이 재정건전성에 미치는 영향을 탐색하고자 하였다는 점과 지방소멸 요인이 재정적 요인에 영향을 줄 수 있는지에 대한 요인을 확인하였다는 점에서 기여를 찾을 수 있다. 특히, 통상적으로 지방소멸 위험이 커지면 재정건전성에 악화를 가져올 수 있다는 현재의 통념과 대비되게 재정적인 측면에서는 현재 이전재원의 역할 등으로 지방소멸 위험과 반비례 관계를 가질 수 있음을 밝힌 점 등이 기여라고 할 수 있다.

참고문헌

- 강경표·김준기. (2020). 우리나라 지방의회의 예산권한이 재정건전성에 미친 영향. 「행정논총」, 58(2): 1-33.
- 강남호. (2011). 지방재정의 건전성 제고 방안. 「산업경제연구」, 24(1): 353-369.
- 고문의·김결. (2021). 한국 지방소멸위험의 공간 분포 변화 분석. 「한국지도학회지」, 21(1): 65-74.
- 김동현·박형준·남승하. (2013). 지방공기업과 지방자치단체의 재정건전성 관계에 관한 연구. 「한국정책과학학회보」, 17(1): 89-117.
- 라영재·정원희·김예승. (2012). 「경기도 재정준칙 도입방안」. 경기도연구용역보고서.
- 박덕병·윤유식·이민수. (2007). 농촌관광마을 지역주민의 사회자본 결정요인: 다중로지스틱 회귀 분석. 「관광레저연구」, 19(4): 27-46.
- 신선정·정원균·안용순·마득상·박덕영·정세환. (2011). 다중로지스틱 회귀모형을 이용한 구강건강 관련 삶의 질과 사회경제적위치의 연관성 분석. 「대한구강보건학회지」, 35(3): 297-305.
- 안국찬. (2013). 지방재정의 건전성 제고방안. 「한국자치행정학보」, 27(4): 41-56.
- 오영호·편도창. (2012). 지방자치단체의 재정특성이 지방채무에 미치는 영향에 관한 연구. 「경영교육연구」, 27(3): 485-515.
- 윤석완. (2010). 시지역 인구 및 고령화와 재정수지의 관계분석. 「재정정책논집」, 12(2): 53-74.
- 위정복. (2015). 지방재정 수입·지출의 구조가 재정운용이 재무적 건전성에 미치는 영향. 「한국공공관리학보」, 29(3): 177-205.
- 이성근·고수정·서준교. (2016). 지방재정의 건전성에 영향을 미치는 요인에 관한 연구. 「지방정부연구」, 19(4):25-42.
- 이진수·김재훈·최승필·임현. (2020). 「재정건전성의 법적 개념과 기준」. 한국법제연구원.
- 이현정·조정래. (2019). 지방자치단체 재정건전성 결정요인: 지방세 비과세·감면이 지방자치단체 재정수지와 부채수준에 미치는 영향을 중심으로. 「한국지방자치학회보」, 31(3): 91-117.
- 임성일·서정섭. (2004). 지방재정분석 및 평가지표 실용성 검증. 「한국지방재정논집」, 9(2): 139-159.
- 최예나·김상현. (2017). 지역특성이 지방재정 건전성에 미치는 영향에 관한 연구: 경기지역 기초

- 자치단체들에 대한 퍼지셋분석을 중심으로. 『한국거버넌스학회보』, 24(1): 155-182.
- 홍영교·최남희. (2021). 지역의 인구구조 변화에 따른 지방재정의 동태성: 시스템사고에 따른 상호작용의 피드백구조 분석. 『미래연구』, 6(1): 161-197.
- Bayaga, A. (2010). Multinomial Logistic Regression: Usage and Application in Risk Analysis. *Journal of applied quantitative methods*, 5(2).
- Berne Robert and Richard Schramm. (1986). *The Financial Analysis of Governments*. Prentice-Hall.
- Beth Walter Honadle, Beverly Cigler and James M. Costa. (2003), *Fiscal Health for Local Government*. Elsevier academic press.
- Böhning, D. (1992). Multinomial logistic regression algorithm. *Annals of the institute of Statistical Mathematics*, 44(1): 197-200.
- Breiman, L. (1996). Bagging predictors. *Machine learning*, 24(2): 123-140.
- Breiman, L. (2001). Using iterated bagging to debias regressions. *Machine Learning*, 45(3): 261-277.
- Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Cart. Classification and Regression Trees*.
- Friedman, J. (1999). *Stochastic gradient boosting*. Department of Statistics.
- Guo, G., Wang, H., Bell, D., Bi, Y., and Greer, K. (2003). KNN model-based approach in classification. In *OTM Confederated International Conferences "On the Move to Meaningful Internet Systems"* (pp. 986-996). Springer, Berlin, Heidelberg.
- Kwak, C., & Clayton-Matthews, A. (2002). Multinomial logistic regression. *Nursing research*, 51(6): 404-410.
- Myles, A. J., Feudale, R. N., Liu, Y., Woody, N. A., and Brown, S. D. (2004). An introduction to decision tree modeling. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 18(6): 275-285.
- Noble, W. S. (2006). What is a support vector machine?. *Nature biotechnology*, 24(12): 1565-1567.
- OECD, (2013). "Fiscal sustainability", in *Government at a Glance 2013*. OECD Publishing, Paris.
- Suthaharan, S. (2016). Support vector machine. In *Machine learning models and algorithms for big data classification* (pp. 207-235). Springer, Boston, MA.
- Ward, Robert B. and Lucy Dadayan. (2009), *State and Local Finance: Increasing Focus on Fiscal Sustainability*. *The Journal of Federalism*, 39(3):455-475.

장유미(張裕美): 연세대학교에서 행정학석사를 취득하고 현재 연세대학교 행정학과 박사과정을 수료하였다. 재무행정, 제도분석, 지방재정이 관심분야이다. 주요 논문으로는 <국고보조금 규모 결정요인 분석: 지방자치단체장의 국회 예산결산특별위원회 네트워크를 중심으로>, <국민참여예산제도의 사업채택요인에 관한 연구> 등이 있다.(93dbal@naver.com)

유한별(劉韓別): 연세대학교에서 행정학 석사, 박사학위를 취득하였다. 관심 분야는 정부평가, 지방소멸, 규제정책, 머신러닝 등이다. 주요논문으로는 <한국 지방소멸 요인과 극복 방안에 관한 연구: 머신러닝 방법을 통한 탐색>, <규제 정책 도입 과정에서 민간 이해관계자의 행태에 관한 연구: 규제혁신형 플랫폼 택시 도입에서 업계 간 갈등을 중심으로> 등이 있다.(yhb5898@gmail.com)

<논문접수일: 2023. 1. 17 / 심사개시일: 2023. 1. 19 / 심사완료일: 2023. 2. 16>

부록

1. 데이터 기술통계

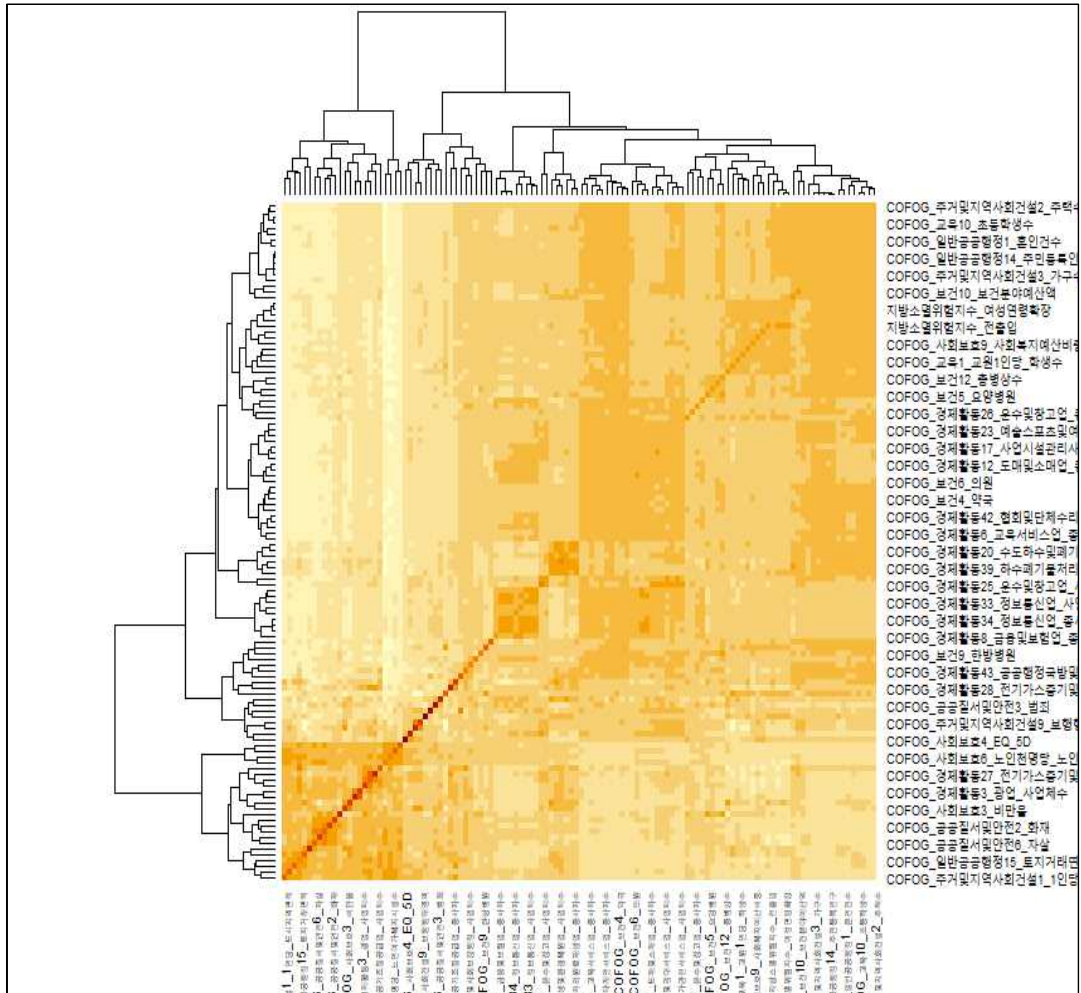
| 구분 | vars | n | mean | sd | min | max |
|-------------------------------------|------|------|----------|----------|-------|--------|
| ID | 1 | 1435 | 37367.78 | 12008.04 | 11110 | 48890 |
| 년도 | 3 | 1435 | 2018 | 2.000697 | 2015 | 2021 |
| 지방소멸위험지수_일반 | 5 | 1435 | 0.711345 | 0.45251 | 0.11 | 2.27 |
| COFOG_경제활동1_건설업_사업체수 | 9 | 1435 | 675.0244 | 847.1404 | 43 | 10619 |
| COFOG_경제활동10_농업임업및어업_종사자수 | 10 | 1435 | 174.528 | 216.9517 | 0 | 1886 |
| COFOG_경제활동11_도매및소매업_사업체수 | 11 | 1435 | 2616.203 | 3432.806 | 47 | 36793 |
| COFOG_경제활동12_도매및소매업_종사자수 | 12 | 1435 | 10681.03 | 15051.93 | 140 | 165267 |
| COFOG_경제활동13_보건업및사회복지서비스업_사업체수 | 13 | 1435 | 538.9172 | 484.6922 | 34 | 3467 |
| COFOG_경제활동14_보건업및사회복지서비스업_종사자수 | 14 | 1435 | 7091.215 | 6434.528 | 342 | 47305 |
| COFOG_경제활동15_부동산업및임대업_사업체수 | 15 | 1435 | 441.2617 | 725.82 | 0 | 11363 |
| COFOG_경제활동16_부동산업및임대업_종사자수 | 16 | 1435 | 1889.126 | 3146.957 | 0 | 42946 |
| COFOG_경제활동17_사업시설관리사업지원및임대서비스업_사업체수 | 17 | 1435 | 254.5889 | 348.9171 | 4 | 3749 |
| COFOG_경제활동18_사업시설관리사업지원및임대서비스업_종사자수 | 18 | 1435 | 4738.673 | 9301.973 | 11 | 86539 |
| COFOG_경제활동19_수도하수및폐기물처리원료재생업_사업체수 | 19 | 1435 | 33.6542 | 36.89544 | 2 | 457 |
| COFOG_경제활동2_건설업_종사자수 | 20 | 1435 | 6080.141 | 8150.385 | 270 | 75538 |
| COFOG_경제활동20_수도하수및폐기물처리원료재생업_종사자수 | 21 | 1435 | 388.9624 | 429.8643 | 14 | 3850 |
| COFOG_경제활동21_숙박및음식점업_사업체수 | 22 | 1435 | 1775.579 | 1737.532 | 55 | 12776 |
| COFOG_경제활동22_숙박및음식점업_종사자수 | 23 | 1435 | 6809.662 | 7088.069 | 158 | 58385 |
| COFOG_경제활동23_예술스포츠및여가관련서비스업_사업체수 | 24 | 1435 | 250.4625 | 296.4446 | 6 | 2901 |
| COFOG_경제활동24_예술스포츠및여가관련서비스업_종사자수 | 25 | 1435 | 1235.38 | 1265.856 | 17 | 10059 |
| COFOG_경제활동25_운수및창고업_사업체수 | 26 | 1435 | 567.7918 | 1143.608 | 2 | 8558 |
| COFOG_경제활동26_운수및창고업_종사자수 | 27 | 1435 | 3528.231 | 4900.1 | 8 | 52507 |
| COFOG_경제활동27_전기가스증기및공기조절공급업_사업체수 | 28 | 1435 | 64.35627 | 206.6938 | 0 | 2993 |
| COFOG_경제활동28_전기가스증기및공기조절공급업_종사자수 | 29 | 1435 | 328.0037 | 492.3315 | 0 | 4086 |
| COFOG_경제활동29_전문과학및기술서비스업_사업체수 | 30 | 1435 | 431.4049 | 988.7275 | 10 | 17293 |
| COFOG_경제활동3_광업_사업체수 | 31 | 1435 | 4.406237 | 6.30101 | 0 | 83 |
| COFOG_경제활동30_전문과학및기술서비스업_종사자수 | 32 | 1435 | 4115.651 | 10944.93 | 41 | 138557 |

| | | | | | | |
|-------------------------------------|----|------|----------|----------|-------|--------|
| COFOG_경제활동31_전산업_사업체수 | 33 | 1435 | 23703.41 | 20796.48 | 1766 | 115054 |
| COFOG_경제활동32_전산업_종사자수 | 34 | 1435 | 98010.55 | 106073.7 | 6356 | 838446 |
| COFOG_경제활동33_정보통신업_사업체수 | 35 | 1435 | 202.5671 | 575.6684 | 2 | 9237 |
| COFOG_경제활동34_정보통신업_종사자수 | 36 | 1435 | 2327.287 | 7455.925 | 6 | 85746 |
| COFOG_경제활동35_제조업_사업체수 | 37 | 1435 | 1395.022 | 2211.16 | 32 | 27258 |
| COFOG_경제활동36_제조업_종사자수 | 38 | 1435 | 16004.1 | 25580.36 | 189 | 255327 |
| COFOG_경제활동37_출판영상방송통신및정보서비스업_사업체수 | 39 | 1435 | 135.5217 | 415.8969 | 3 | 4290 |
| COFOG_경제활동38_출판영상방송통신및정보서비스업_종사자수 | 40 | 1435 | 2039.97 | 6905.973 | 6 | 64204 |
| COFOG_경제활동39_하수폐기물처리원료재생및환경복원업_사업체수 | 41 | 1435 | 26.74181 | 27.88212 | 2 | 242 |
| COFOG_경제활동4_광업_종사자수 | 42 | 1435 | 58.81316 | 152.1248 | 0 | 1853 |
| COFOG_경제활동40_하수폐기물처리원료재생및환경복원업_종사자수 | 43 | 1435 | 305.7768 | 337.4117 | 4 | 2290 |
| COFOG_경제활동41_협회및단체수리및기타개인서비스업_사업체수 | 44 | 1435 | 831.6106 | 855.0849 | 66 | 7524 |
| COFOG_경제활동42_협회및단체수리및기타개인서비스업_종사자수 | 45 | 1435 | 2642.87 | 2687.176 | 118 | 24708 |
| COFOG_경제활동43_공공행정국방및사회보장행정_사업체수 | 46 | 1435 | 50.57073 | 19.40858 | 10 | 127 |
| COFOG_경제활동44_공공행정국방및사회보장행정_종사자수 | 47 | 1435 | 3205.507 | 2611.362 | 692 | 20775 |
| COFOG_경제활동5_교육서비스업_사업체수 | 48 | 1435 | 446.2397 | 542.5138 | 6 | 5285 |
| COFOG_경제활동6_교육서비스업_종사자수 | 49 | 1435 | 3914.594 | 4505.974 | 41 | 33740 |
| COFOG_경제활동7_금융및보험업_사업체수 | 50 | 1435 | 175.2169 | 281.7821 | 13 | 4981 |
| COFOG_경제활동8_금융및보험업_종사자수 | 51 | 1435 | 2958.178 | 6619.462 | 109 | 67009 |
| COFOG_경제활동9_농업임업및어업_사업체수 | 52 | 1435 | 22.10819 | 28.03064 | 0 | 329 |
| COFOG_공공질서및안전1_교통사고 | 53 | 1435 | 3.075171 | 1.029793 | 1 | 5 |
| COFOG_공공질서및안전2_화재 | 54 | 1435 | 3.068028 | 1.01726 | 1 | 5 |
| COFOG_공공질서및안전3_범죄 | 55 | 1435 | 2.985582 | 1.101696 | 1 | 5 |
| COFOG_공공질서및안전5_생활안전 | 56 | 1435 | 3.064843 | 1.093451 | 1 | 5 |
| COFOG_공공질서및안전6_자살 | 57 | 1435 | 3.067331 | 1.046254 | 1 | 5 |
| COFOG_공공질서및안전7_감염병 | 58 | 1435 | 3.065631 | 1.064883 | 1 | 5 |
| COFOG_교육1_교원1인당_학생수 | 59 | 1435 | 13.31922 | 4.435256 | 4.77 | 26.78 |
| COFOG_교육10_초등학생수 | 60 | 1435 | 10230.29 | 10093.45 | 451 | 68101 |
| COFOG_교육11_유치원학급당_학생수 | 61 | 1435 | 16.07538 | 4.082795 | 5 | 25.55 |
| COFOG_교육12_초등학교학급당_학생수 | 62 | 1435 | 18.69091 | 4.985238 | 8.03 | 27.91 |
| COFOG_교육13_중학교학급당_학생수 | 63 | 1435 | 23.12347 | 4.752195 | 10.37 | 33.17 |
| COFOG_교육14_고등학교학급당_학생수 | 64 | 1435 | 24.28199 | 4.356731 | 10.89 | 35.36 |
| COFOG_교육2_재적학생수 | 65 | 1435 | 34177.39 | 34092.07 | 1045 | 178513 |
| COFOG_교육3_교원수 | 66 | 1435 | 2172.108 | 1862.564 | 195 | 10094 |
| COFOG_교육4_유치원교원수 | 67 | 1435 | 201.9983 | 200.0904 | 9 | 1457 |
| COFOG_교육5_유치원수 | 68 | 1435 | 34.31185 | 25.51185 | 3 | 180 |

| | | | | | | |
|---------------------------------|-----|------|----------|----------|-------|---------|
| COFOG_교육7_천명당_사설학원수 | 69 | 1435 | 1.259972 | 0.599504 | 0.2 | 7.4 |
| COFOG_교육8_사설학원수 | 70 | 1435 | 298.8167 | 331.5926 | 6 | 2383 |
| COFOG_교육9_초등교원수 | 71 | 1435 | 721.639 | 599.4677 | 76 | 4111 |
| COFOG_보건1_병원 | 72 | 1435 | 5.826132 | 6.17232 | 0 | 34 |
| COFOG_보건10_보건분야예산액 | 73 | 1435 | 12702.16 | 6541.584 | 2345 | 62004 |
| COFOG_보건11_천명당_의료기관병상수 | 74 | 1435 | 15.02553 | 9.760323 | 0 | 70.1 |
| COFOG_보건12_총병상수 | 75 | 1435 | 2731.798 | 2422.163 | 0 | 12351 |
| COFOG_보건2_보건소 | 76 | 1435 | 0.962927 | 0.330196 | 0 | 3 |
| COFOG_보건3_상급종합병원 | 77 | 1435 | 0.163415 | 0.420524 | 0 | 2 |
| COFOG_보건4_약국 | 78 | 1435 | 87.75383 | 78.40979 | 3 | 486 |
| COFOG_보건5_요양병원 | 79 | 1435 | 5.839721 | 5.537978 | 0 | 36 |
| COFOG_보건6_의원 | 80 | 1435 | 124.1993 | 156.4117 | 1 | 1752 |
| COFOG_보건7_종합병원 | 81 | 1435 | 1.212892 | 1.385749 | 0 | 7 |
| COFOG_보건8_치과병원 | 82 | 1435 | 0.944948 | 1.790092 | 0 | 17 |
| COFOG_보건9_한방병원 | 83 | 1435 | 1.330314 | 3.401783 | 0 | 40 |
| COFOG_사회보호1_건강보험_적용인구_현황 | 84 | 1435 | 194767.7 | 170249.8 | 15822 | 901163 |
| COFOG_사회보호10_사회복지분야예산액 | 85 | 1435 | 200688.6 | 145077.4 | 30715 | 1110065 |
| COFOG_사회보호2_고위험층주울 | 86 | 1435 | 18.11307 | 4.043284 | 4.9 | 31.4 |
| COFOG_사회보호3_비만율 | 87 | 1435 | 29.29206 | 3.505889 | 18.4 | 40.7 |
| COFOG_사회보호4_EQ_5D | 88 | 1435 | 0.956955 | 0.009492 | 0.93 | 0.99 |
| COFOG_사회보호5_주관적_건강수준인지율 | 89 | 1435 | 46.81746 | 7.477105 | 29.5 | 71.6 |
| COFOG_사회보호6_노인천명당_노인여가복지시설수 | 90 | 1435 | 10.02955 | 7.645895 | 1.2 | 32.7 |
| COFOG_사회보호7_노인여가복지시설수 | 91 | 1435 | 285.1814 | 156.0109 | 30 | 828 |
| COFOG_사회보호8_주민등록인구60세이상 | 92 | 1435 | 33128.45 | 22093.94 | 4493 | 118668 |
| COFOG_사회보호9_사회복지예산비중 | 93 | 1435 | 34.08523 | 15.19155 | 13.2 | 71.5 |
| COFOG_일반공공행정1_혼인건수 | 94 | 1435 | 934.0261 | 931.6307 | 34 | 4982 |
| COFOG_일반공공행정10_전년총인구 | 95 | 1435 | 203857.7 | 176216.8 | 17696 | 874745 |
| COFOG_일반공공행정11_일반공공행정예산비중 | 96 | 1435 | 7.192014 | 3.417637 | 2.04 | 32.49 |
| COFOG_일반공공행정12_일반공공행정분야예산액 | 97 | 1435 | 61206.07 | 45621.02 | 10535 | 424671 |
| COFOG_일반공공행정13_천명당_외국인수 | 98 | 1435 | 22.0639 | 17.38566 | 2.5 | 103.85 |
| COFOG_일반공공행정14_주민등록인구 | 99 | 1435 | 199659.7 | 174687.3 | 16320 | 887015 |
| COFOG_일반공공행정15_토지거래면적 | 100 | 1435 | 8272.003 | 7150.471 | 5 | 37197 |
| COFOG_일반공공행정2_인구십만명당_자살률 | 101 | 1435 | 29.20808 | 7.787405 | 7.8 | 70.6 |
| COFOG_일반공공행정3_합계출산율 | 102 | 1435 | 1.113763 | 0.300206 | 0.38 | 2.54 |
| COFOG_일반공공행정4_조혼인율 | 103 | 1435 | 4.737645 | 1.154805 | 2.2 | 9.1 |
| COFOG_일반공공행정5_출생아수 | 104 | 1435 | 1265.252 | 1285.084 | 40 | 7262 |
| COFOG_일반공공행정6_평균연령 | 105 | 1435 | 45.17822 | 4.91305 | 35 | 58.3 |
| COFOG_일반공공행정7_남성성비 | 106 | 1435 | 100.1753 | 5.126333 | 88.79 | 128.1 |
| COFOG_일반공공행정8_인구증가율 | 107 | 1435 | 1.38724 | 1.718384 | -4.31 | 26.4 |
| COFOG_일반공공행정9_당해년총인구 | 108 | 1435 | 204224.2 | 178760.6 | 16575 | 922092 |
| COFOG_주거및지역사회건설1_1인당_도시지역 면적 | 109 | 1435 | 596.7118 | 595.8574 | 35.6 | 4736.54 |

| | | | | | | |
|------------------------------|-----|------|----------|----------|-------|----------|
| COFOG_주거및지역사회건설10_인당_자동차등록대수 | 110 | 1435 | 0.49053 | 0.141452 | 0.2 | 2 |
| COFOG_주거및지역사회건설2_주택수 | 111 | 1435 | 69164.58 | 55812.18 | 7445 | 311842 |
| COFOG_주거및지역사회건설3_가구수 | 112 | 1435 | 80065.49 | 68429.01 | 7609 | 356838 |
| COFOG_주거및지역사회건설4_하수도보급률 | 113 | 1435 | 84.65398 | 16.26632 | 30.9 | 100 |
| COFOG_주거및지역사회건설5_상수도보급률 | 114 | 1435 | 90.93923 | 12.0621 | 44.1 | 100 |
| COFOG_주거및지역사회건설6_교통문화지수 | 115 | 1435 | 78.66417 | 5.579159 | 55.38 | 92.61 |
| COFOG_주거및지역사회건설7_운전행태영역 | 116 | 1435 | 42.67813 | 5.863254 | 3.2 | 53.56 |
| COFOG_주거및지역사회건설8_교통안전영역 | 117 | 1435 | 19.37898 | 6.790832 | 3.98 | 38.56 |
| COFOG_주거및지역사회건설9_보행행태영역 | 118 | 1435 | 14.96806 | 2.251859 | 3.75 | 19.06 |
| COFOG_환경보호1_녹지지역면적 | 119 | 1435 | 47723518 | 61705763 | 25135 | 4.11E+08 |
| COFOG_환경보호2_녹지지역면적비율 | 120 | 1435 | 64.74166 | 20.91531 | 0.25 | 92.97 |
| COFOG_휴양문화종교1_십만명당_문화기반시설수 | 121 | 1435 | 10.58873 | 9.428897 | 1 | 74.9 |
| COFOG_휴양문화종교2_문화기반시설수 | 122 | 1435 | 11.19748 | 6.866853 | 2 | 66 |
| 통합재정수지비율 | 123 | 1435 | 0.056279 | 0.059421 | -0.18 | 0.5 |
| 예산대비채무비율(btI 포함) | 124 | 1435 | 0.019387 | 0.034094 | 0 | 0.34 |
| 채무상환비율 | 125 | 1435 | 0.006527 | 0.009014 | 0 | 0.062966 |

2. 특성변수(feature) 상관관계 분석 히트맵(heatmap)



※ 상기 제시한 히트맵은 독립변수 간 상관관계를 시각화 한 자료이다. 기존 회귀분석을 사용한 연구에서는 상관관계 분석 후 다중공선성을 고려하여 상관성을 보이는 변수를 제거한 후 모형을 구축하나, 본 연구에서는 머신러닝의 장점을 활용하기 위해 모형에 모든 변수를 반영하여 수행하였다.

Abstract

Exploring Factors Influencing Local Government's Fiscal Health Using Machine Learning

Jang, Youmi
Yoo, Hanbyeol

This study focused on the demographic structure and socioeconomic environmental factors to confirm the impact on the financial health of local governments. As a result, it was confirmed that social and economic environmental factors had a major effect on the financial health of local governments. Specifically, the welfare factor and the economic factor had an major effect. In the welfare factor, it was confirmed that the proportion of the social welfare budget had a negative(-) effect on the debt repayment ratio and the debt-to-budget ratio. In the case of the economic factor, the influence of each industry was confirmed through machine learning. Although the results were not statistically significant, the primary industries such as agriculture and fishing had a positive(+) effect on the balance of payments, while secondary and tertiary industries had a negative(-) effect on the balance of payments. The study confirmed that there is the dependence of local governments on the central government's transfer resources.

Key Words: local government, local finance, fiscal health, local extinction, machine learning