

기계학습과 네트워크를 이용한 수학교육 연구의 영향력 판별 프로그램 개발

오 세 준 (이화여자대학교 사범대학부속 이화·금란고등학교, 교사)
권 오 남 (서울대학교, 교수)[†]

본 연구는 수학교육에서 영향력 있는 논문을 판별하는 기계학습 프로그램 개발 연구이다. 이를 위하여 과학계량학의 관점에서 논문의 영향력을 조명하고, 수학교육 연구 네트워크를 구성하고, 네트워크 중심성 지수인 PageRank로 수학교육 연구의 영향력으로 정의하였다. 영향력 있는 수학교육 연구를 판별하기 위하여 기계학습 모델을 설계하였으며, 이를 이용하여 영향력 있는 논문이 게재된 비율이 높은 학술지를 조사한 결과 Journal for Research in Mathematics Education(25.66%), Educational Studies in Mathematics(22.12%), Zentralblatt für Didaktik der Mathematik(8.46%), Journal of Mathematics Teacher Education(5.8%), Journal of Mathematical Behavior(5.51%) 순으로 나타났다. 수학교육 전문가들이 직접 논문을 읽고 질적으로 평가한 선행연구 결과와 유사한 결과를 기계학습 프로그램으로 도출할 수 있었다. 많은 인원과 시간이 필요했던 수학교육 연구의 영향력 평가를 인공지능을 이용하여 효율적으로 실시할 수 있었다는 점에서 의의가 있다.

I. 서론

수학교육 연구는 지난 100년간 수학 교수·학습 과정을 규명하면서 독자적인 학문으로서 발전하였다. 즉, '수학'과 '교육'이라는 단어가 단순하게 결합된 용어가 아닌, '수학교육'의 정체성을 확보하며 독립적인 연구 분야로 발전하는 과정이 있었다. 수학교육 연구 분야의 발전은 수학교육 연구들이 누적되고, 누적된 연구성과를 분석하며 미래 수학교육 연구를 위한 방향을 제시하는 노력이 있었기에 가능하였다.

한편 매일 전 세계의 수학교육 연구자들은 자신의 연구 질문과 관련된 유용한 논문을 찾아내기 위하여 많은 양의 수학교육 연구 논문을 찾고 읽는다. 연구자는 연구의 필요성을 제시하기 위하여 자신의 연구와 선행연구를 비교하고, 연구결과를 정당화하기 위하여 다른 사람의 연구결과를 참조하고 인용한다. 이때, 수학교육 연구에 대한 메타분석 연구들은 좋은 지침서가 될 수 있다. 하지만 수학교육 연구자들이 확인할 수 있는 메타분석 연구는 논문 집필에서 게재확정까지 최소 1~2년의 세월이 필요하기에 가장 최신의 수학교육 연구를 담지 못한다. 또한, 지금까지 컴퓨터를 활용한 메타분석 연구들은 수학교육 연구 분야의 토픽을 분류할 수 있지만, 개별 논문의 영향력은 확인할 수 없다.

이와 같은 필요성으로 수학교육 메타분석 연구들이 있었다. National Council of Teachers of Mathematics(이하 NCTM)는 『수학 교수와 학습에 관한 연구 핸드북』(Grouws, 1992)을 발간하면서 “과거 수학교육 연구의 통합과 개념화 그리고 수학교육 분야를 발전시키는 데 가장 유용한 연구를 제안”을 핸드북 발간의 목적으로 하

* 접수일(2023년 2월 24일), 심사(수정)일(2023년 3월 10일), 게재확정일(2023년 3월 14일)

* MSC2000분류 : 97B10

* 주제어 : 수학교육 연구, 네트워크, 영향력, 기계학습

[†] 교신저자 : onkwon@snu.ac.kr

* 이 논문은 오세준의 박사학위 논문 「기계학습과 네트워크를 이용한 수학교육 연구의 메타분석」의 일부를 수정 보완하였음.

였다. 이후 NCTM은 최신 수학교육 연구를 지속적으로 소개하기 위하여 두 번째 개정판과 세 번째 개정판을 출간하기도 하였다. 마찬가지로 English, Kirshner(2002)는 『국제 수학교육 핸드북』을 통해 수학교육의 연구성과를 확인하고 수학교육 연구의 발전적 방향을 제시하였다. 『국제 수학교육 핸드북』도 현재까지 3번의 개정판을 출간하며, 최신 수학교육 연구를 소개하고 있다. 한편 국내 수학교육 연구 동향을 분석하고 국내 수학교육 연구자들에게 발전 방향과 후속연구에 대한 아이디어를 제시한 연구(박경미, 2003; 박선영, 김원경, 2011; 방정숙 외, 2019)도 있었으며, 국내외 수학교육 연구를 양적으로 분석하고 질적으로 비교한 연구(권오남, 주미경, 2003)도 있었다.

수학교육 연구자들은 관심 분야의 논문을 직접 검색하여 읽고 정리해야 한다. 하지만 인용하기에 적절하지 않은 연구들도 있으며(Wiliam, 2003), 때때로 연구결과가 모호하거나 서로 모순되는 결과를 낳는 예(Johnson, 1987)도 있다. 이에 수학교육 연구의 영향력을 평가하는 것은 의미 있다. Forgasz(2019)는 수학교육 분야의 신진 연구자들을 위한 개요서에서 연구자가 인용하는 논문은 평판이 좋아야 한다고 주장하였다. 신진연구자들이 영향력이 높은 논문을 쉽게 판별할 수 있고, 이를 바탕으로 영향력 있는 논문을 쓸 수 있다면 수학교육 연구 분야의 발전이 가능할 것이다.

물론 수학교육 연구의 구성요소를 정의하는 과정에서 수학교육 연구의 영향력을 평가하기 위한 노력이 있었다(Sierpiska & Kilpatrick, 1998). 수학교육 연구의 영향력을 평가하는 중요한 요소 중 하나는 연구가 다른 연구자에게 얼마나 유용한지에 대한 수학교육 연구자들의 의견이다(Andrade-Molina et al., 2020; Lester, 2005; Schoenfeld, 2007). 다른 연구자에게 얼마나 유용한지는 피인용횟수를 활용하여 분석할 수 있다(Garfield & Sher, 1963). 성공적인 출판물은 인용될 가능성이 더 크고, 생산적인 저자는 논문을 다시 출판할 가능성이 더 크며, 많은 연구자가 일반적으로 참고하는 학술지는 인용될 가능성이 더 크기 때문이다(Price, 1976).

수학교육 연구의 영향력을 평가하는 또 다른 중요한 요소 중 하나는 연구가 연구의 새로움(novelty), 타당성(validity) 및 중요성에 대한 수학교육 연구자들의 의견이다. 연구의 새로움, 타당성, 중요성은 해당 분야의 지식 있는 전문가의 의견을 구함으로써 각각 측정할 수 있다. 그러나 과학 출판물의 수는 기하급수적으로 증가(Price, 1963)하고 있으므로, 전문가들이 출판된 수학교육 연구 논문을 평가하는 것은 많은 시간이 필요하다. 이에 수학교육 연구의 영향력 평가대상은 개별 논문이 아닌 학술지였다(Williams & Leatham, 2017). 학술지에 게재된 논문은 이미 새로움, 타당성, 중요성은 검증되었다고 믿는 것이다.

하나의 학술지가 다른 학술지보다 더 자주 인용되거나 질적으로 우수한 논문을 게재하는 것으로 평가된다면, 그 학술지는 높은 평가를 받는다. 또한, 그 학술지에 게재된 논문도 높은 평가를 받는 것으로 간주하는 방식이었다. 하지만 학술지의 영향력 지수가 개별 논문의 영향력을 대변하는 것은 아니며(Seglen, 1997), 학술지를 대상으로 하는 어떤 인용 지표를 개별 논문의 중요성을 평가하기 위한 용도로 쓰여서는 안 된다(Neylon & Wu, 2009)는 의견도 있다. 그러나 수학교육 논문을 하나씩 평가하는 것은 또다시 많은 인적 자원과 시간이 필요하다.

따라서 수학교육 연구의 영향력을 파악하기 위한 새로운 대안이 필요하며, 그중 하나로 인공지능 기술을 활용하는 방안이 제시되고 있다. 빅데이터와 인공지능을 활용하면 논문의 영향력을 평가하고 예측 가능함(Weihs & Etzioni 2017; Zhou et al., 2021; Zhu et al., 2015)이 연구자들에 의해 밝혀져 왔다. 특히 단순히 피인용횟수가 아닌 논문-저자-학술지의 네트워크를 구성하면 논문의 영향력을 더 효과적으로 예측(Mariani et al., 2016; Weis & Jacobson, 2021)할 수 있기에 인공지능을 활용하여 논문의 영향력을 예측하는 방법은 주목받고 있으며 수학교육 분야에서도 이에 대한 논의가 요구된다.

이에 본 연구는 인공지능을 활용하여 수학교육 연구에 대한 메타분석을 실시하고자 한다. 이 연구의 목적은 수학교육에서 영향력 있는 논문을 판별하여 수학교육 연구에 발전적인 시사점을 제공하고자 하였다. 먼저 과학계량학의 관점에서 영향력 있는 수학교육 연구를 판별하기 위하여 기계학습으로 영향력 높은 논문을 판별하는 모델을 설계하고자 한다. 구체적인 연구 질문은 다음과 같다.

수학교육의 영향력 높은 논문은 어떻게 판별할 수 있을까?

II. 연구의 배경

1. 이론적 배경

가. 과학계량학

과학계량학은 과학(지식)이 어떻게 구조화되고 생성되며 상호 연관되는지에 대한 정량적 연구(Hood & Wilson, 2001)이며, 여기에는 과학 주변의 구체적이면서도 커뮤니케이션 구성요소가 포함된다(Wilson, 2001). Nalimov, Mulchenko(1971)는 "정보화 과정으로서의 과학 발전에 관한 연구의 정량적 방법"을 설명하기 위해 '과학계량학(Scientometrics)'이라는 용어를 제안하였다. 이후 Price(1976)에 의해 과학 논문을 평가하는 정량적 기법과 그것들을 모델링 하는 방법이 확립되었으며, 과학적 방법을 사용하여 과학을 연구하기 위한 기초를 만들었다. 이에 과학계량학은 때때로 "과학의 과학(science of science)"이라고 불린다. 특히 Price는 사회 과학의 다양한 분야에 폭넓게 적용될 수 있는 통계의 확률분포를 기반으로 하는 누적 우위 분포(cumulative advantage distribution)의 개념을 과학계량학에 도입하였다. 연구자가 인용하는 문헌은 소수의 학술지에 집중적으로 게재되며, 상대적으로 소수의 단어가 반복되는 등 지금까지 경험적으로 알고 있는 사실들을 수학적 이론으로 설명하기 위한 기본적인 방법론(Lotka의 분포, Bradford의 법칙 등)을 확립하였다. Price가 도입한 개념과 이론은 과학 출판물을 분석하는 과학계량학의 근간이 되었다(Garfield, 2009).

지식의 발전을 이해하는 것은 현재의 의사결정, 미래 예측 등에 영향을 미치기 때문에 중요하다. 과학은 누적되며, 이러한 누적된 지식은 후속 연구자들의 발전에 도움이 된다(Garfield, 2009). 한편 지식 발전은 공공과 기업의 이익과도 직결된다(Altuntas et al., 2015; Bildosola et al., 2017). 기업 및 공공분야에서 중요한 의사결정을 하기 전에, 의사결정자는 지식에 대한 평가와 예측을 수행해야 한다. 하지만 현재 이용 가능한 제한된 정보로는 미래를 예측하기가 쉽지 않기 때문에, 지식의 진화에 관한 연구는 수십 년의 노력에도 불구하고 완전히 달성되지 못했다.

지식의 진화를 예측하는 것은 정보, 기술, 과학을 과학적으로 조사할 수 있는 대상으로 간주함으로써 수집된 지식의 데이터에서 패턴을 판별하고 향후 발전을 예측하는 것으로 구성된다. 과학계량학에서 지식의 진화과정을 예측하는 연구는 두 가지 방법으로 진행되고 있다(Weihs & Etzioni, 2017). 첫 번째 연구 방법은 피인용횟수를 통계적으로 모델링하는 Price(1963) 연구의 정신적 계승이다. 예를 들어, Price(1976)는 인용 네트워크의 모델을 구성하기 위해 과학 출판물을 집계했고, 한 출판물은 확률적으로 다른 출판물에 연결되었다고 주장하였다. Price는 과학 출판물의 분포는 치우친 분포나 쌍곡선 분포와 유사할 수 있으며 영향력 있는 논문들을 바탕으로 더 큰 영향력 있는 논문이 탄생하는 과정을 설명하는 누적 우위 분포를 도입하며, 확률적으로 지식의 진화과정을 예측하려고 시도하였다. 두 번째 연구 방법은 인공지능을 활용하여 논문의 영향력을 예측하는 것이다. 논문이 가진 여러 특성을 메타데이터로 입력하면 기계학습 혹은 딥러닝을 활용하여 영향력 있는 논문을 예측할 수 있다. 또한, 인공지능을 활용한 방법은 확률적 분포를 활용하여 예측하는 방법이 지닌 정적인 한계를 극복할 수 있다(Acuna et al., 2012; Dong et al., 2015).

두 가지 방법론의 기본 가정은 하나의 논문은 독립적으로 존재하지 않고, 연구 네트워크 속에 함께 있다는 것이다. 즉, 학술 연구에서 논문은 단독으로 존재하는 것이 아니라, 그 학문 분야의 여러 논문을 근거로 지탱되고 있으며 이때, 인용은 논문과 다른 논문 사이의 관계를 보여주는 수단이다. Merton(1973)은 "과학의 체계에서 타인에 의해 주목되고 사용되지 않으면 그 가치에 대한 의심이 높아지기 쉽다"라는 주장과 함께 인용의 중요성을 강조하였다. 다시 말해 연구 네트워크에서 인용 관계를 분석하는 것은 하나의 논문이 연구 분야에 미치는 영향을 결정하는 것이다(Zhu et al., 2015). 따라서 이 연구에서도 영향력이 높은 논문을 판별하기 위하여, 인용 관

계를 중심으로 수학교육 네트워크를 형성하고 분석할 것이다.

나. 네트워크와 논문의 영향력 지수

수학교육 분야에서도 연구의 네트워크는 중요하다. 1988년 NCTM은 “중요한 질문에 관한 연구 방향”과 “협업 네트워크를 지원하기 위한 메커니즘 개발 장려”를 위해 『수학교육 연구 과제』(Grouws, 1988)를 발간하며 협업 네트워크를 장려하였다.

수학교육 지식은 결국 논문으로 확인될 수 있으며, 논문과 논문은 인용을 매개로 네트워크를 형성한다. Fujigaki(1998)는 과학 논문을 인용하는 것은 다음 시대의 지식개발에 지속해서 사용될 지식을 선택하는 과정이기 때문에 과학 분야의 성장에 중요한 부분이라고 강조하였다. 즉, 인용을 바탕으로 새로운 지식이 탄생하기 때문에 인용은 진화의 매개체이며, 지식의 전달자이다. 따라서 해당 학문 분야의 인용 네트워크는 지식의 진화를 표현한 것이다.

이는 2006년 NCTM이 『수학 교수와 학습에 관한 연구 핸드북』(Grouws, 2006)을 발간한 이유에서도 확인할 수 있다. NCTM은 “과거 수학교육 연구의 통합과 개념화 그리고 수학교육 분야를 발전시키는 데 가장 유용한 연구를 제안”을 핸드북 발간의 목적으로 하였다. 다시 말해 수학교육 분야를 발전시키는 유용한 논문을 제안할 수 있으면, 그 논문은 후속 연구자들에게 인용되면서 수학교육 분야의 발전을 이끌 것이다.

한편 수학교육 분야에서도 유용한 연구와 연결을 찾는 네트워크 이론에 관한 연구가 있었으며, 수학교육 연구자들은 다른 이론적 접근법을 이해하고, 활용하고, 심지어 통합하려고 시도하고 있다(Ingliš & Foster, 2018). Lakatos(1978)는 통합 가능성에 상관없이 서로 다른 연구를 이해하고 참여하는 것이 학문 분야의 발전에 기여할 가능성이 크다고 강조하였다.

하지만 수학교육의 분야의 지식인 논문은 오늘도 계속 생성되고 있기에 개별 연구자가 수학교육 분야의 중요한 논문을 판별하는 것은 많은 시간과 노력이 필요하다. 수학교육 분야를 성장시킬 연구는 어떻게 찾을 수 있을까? 다시 말해 수학교육 네트워크에서 수학교육 분야를 발전시킬 수 있는, 영향력 있는 논문은 어떻게 판별할 수 있을까? 그 답을 과학계량학의 관점으로 찾으려고 한다.

과학계량학에서 연구의 성과를 측정하는 방법으로 생산성과 영향력이 있으며, 생산성은 논문의 수로 측정할 수 있다(김관준, 이재운, 2010). 한편 과학계량학에서 연구의 영향력은 인용을 기반으로 측정되어 왔다. Merton(1973)은 “과학의 체계에서 타인에 의해 주목되고 사용되지 않으면 그 가치에 대한 의심이 높아지기 쉽다”라는 주장과 함께 인용의 중요성을 강조하였다. Zhu 외(2015)는 논문이 연구 분야에 미치는 영향을 결정하는 것은 인용이라고 하였다.

이에 학술지의 영향력을 평가하는 지표는 유진 가필드(Eugene Garfield)의 Impact factor(이하 IF), 개별 논문의 영향력을 평가하는 지표는 Field-Weighted Citation Impact(이하 FWCI)가 활용된다. FWCI는 SCOPUS 등 저널 논문을 대상으로 연구 분야에 대한 상대적인 피인용횟수로 다음과 같은 수식으로 계산할 수 있다.

$$FWCI(A) = \frac{A \text{의 피인용 횟수}}{A \text{와 같은 연도, 주제, 출판물 형태가 같은 논문들의 전 세계 평균 피인용 횟수}}$$

예를 들어 논문 A의 영향력을 나타내는 FWCI(A)의 값이 1이면 논문 A는 전 세계 논문 중 평균에 해당하는 영향력을 지니고 있으며, FWCI(A)의 값이 1.5이면 논문 A는 전 세계 논문들이 받는 평균적인 피인용횟수보다 50% 더 인용되고 있음을 뜻한다.

한편 연구자의 영향력을 평가하는 대표적인 지표로 활용되는 h-index(Hirsch, 2005)도 인용 분포를 기반으로 한다(Bornmann et al., 2008). h-index는 연구자의 전체 논문의 수와 피인용횟수를 바탕으로 산출되는 값이다.

연구자의 전체 논문을 피인용횟수 순으로 내림차순으로 나열한 후, 피인용횟수와 내림차순의 순서를 비교하여 연구자의 h-index를 결정한다. 결국, 연구자의 영향력도 피인용횟수를 기반으로 산출된 것이다.

즉 과학계량학에서 학술지, 연구자, 논문의 영향력은 피인용횟수를 기반으로 측정되고 있다. 피인용횟수를 기반으로 하는 영향력 지수는 자연스럽지만 완전한 지표는 아니다(Chen et al., 2007). 피인용횟수는 몇 가지 한계를 가진다. 첫째, 모든 개별 인용은 균등하게 1로 측정되는데, 모든 인용이 같은 가치를 지니는 것이 아니므로 문헌의 중요도를 정확히 측정한다고 보기 어렵다. 이에 몇몇 연구자들은 연구의 영향력을 측정할 때 인용에 가중치를 두는 것을 제안했다(Aljuaid et al., 2021; Ding & Cronin, 2011; Lee & Yang, 2015). 둘째, 영향력 있는 논문은 오래된 논문일수록 피인용횟수가 높을 가능성이 높아 최근 논문은 불리한 평가를 받을 수 있다. 따라서 영향력 있는 논문을 판별하기 위해서는 피인용횟수뿐 아니라 이를 보완할 수 있는 다른 지표가 필요하다.

과학계량학의 관점에서 논문의 영향력은 피인용횟수뿐 아니라 네트워크 구조도 함께 고려되어 측정된다. 새로운 논문(지식)이 해당 학문 분야의 네트워크에 들어오면 기존 네트워크를 강화하거나 대체하는 방향으로 네트워크 구조에 변화가 일어난다(Funk & Owen-Simth, 2017). 새롭고 영향력 있는 지식은 인용 네트워크 내에서 수많은 연결을 빠르게 생성하는 것으로 나타났다. 다시 말해 네트워크의 구조를 관찰하면 영향력 있는 논문을 판별할 수 있다는 것을 추론할 수 있다.

네트워크의 구조와 관련된 지표들이 피인용횟수를 기반으로 한 지표보다 영향력 있는 논문을 더 잘 판별한다는 것은 여러 연구를 통해서 확인됐으며, 이는 학술 연구 네트워크의 구조에 과학 출판물의 순위를 결정하는 데 사용할 수 있는 정보가 포함되어 있음을 시사한다(Mariani et al., 2016). 피인용횟수만으로 과소평가되어 영향력 없는 논문으로 취급되었지만, 네트워크 구조를 분석하여 숨겨진 보석과 같은 대기만성형 논문을 발견하는 연구(Walker et al., 2007)는 영향력을 평가하는 지표로서 네트워크 구조의 중요성을 보여주는 중요한 사례이다.

네트워크 분석방법은 1967년 조직이론에서 파생되어 조직의 상호작용에 대한 분석에서 시작되었다(조은성, 송재도, 2013). 네트워크 분석방법은 사람들이 관계를 맺으며 연결망(network)을 구축하고, 사회를 형성하는 모습에 착안하였기에 초기에는 사회 연결망으로 번역되어 명명되기도 하였다. 네트워크는 조직이론에 국한되는 것이 아니라 관계를 갖는 모든 영역에 적용될 수 있다. 네트워크 분석은 ‘대상’과 ‘관계’에 주목하여 ‘대상’ 사이의 ‘연결’성을 시각적으로 제시함으로써 추상적인 구조를 구체화한다. 따라서 네트워크 분석으로 ‘대상’과 ‘대상’ 사이의 관계를 파악할 수 있다. 네트워크 분석방법의 핵심은 그래프 이론(Barnes & Harary, 1983)으로 그래프를 구성하는 ‘대상’을 노드(node), 이 ‘대상’들 사이의 ‘연결’ 관계를 엣지(edge)라고 한다.

다. 네트워크 분석방법과 중심성 지수

네트워크 분석방법은 논문이 인용 받은 빈도를 알 수 있게 할 뿐 아니라 이 논문이 다른 논문과 어떤 관계가 있는지, 그 관계의 강도는 얼마인지 파악할 수 있게 한다. 즉 특정 논문이 얼마나 인용되었는지 알 수 있을 뿐만 아니라, 그 논문이 해당 학문 분야에서 어떤 역할과 기여를 하는지, 네트워크 구조를 분석하여 파악할 수 있다.

이에 연구의 영향력을 분석하기 위하여 네트워크 분석방법은 활용되어 왔다. 네트워크 분석방법으로 연구의 영향력을 평가한 연구를 정리하면 <표 II-1>과 같다. 선행연구에서 노드로 활용한 것은 논문, 저자, 학술지였으며, 엣지로 활용한 것은 논문과 논문은 인용 관계, 논문과 저자는 저술 관계, 저자와 저자는 공동저자 관계, 논문과 학술지는 게재 관계, 저자와 소속은 소속 관계 등이 있었다. 본 연구에서 수학교육 네트워크를 구성할 때, Weis, Jacobson(2021)과 같이 노드는 논문, 저자, 학술지, 소속으로 선정하였으며, 엣지는 인용 관계, 공동저자 관계, 저술 관계, 소속 관계, 게재 관계로 구성하였다.

<표 11-1> 영향력 평가를 위하여 네트워크 분석방법을 활용한 선행연구

연도	저자명	노드	엣지	분석방법
2007	이정연 외	논문, 저자	저술 관계	확률적 온톨로지
2008	이재윤	학술지, 논문	게재 관계	공유도 중심성 분석
2011	이재윤	학술지, 논문	게재 관계	중심성 분석 PageRank
2014	권선영	논문, 저자	저술 관계	중심성 분석
2014	이재윤	논문, 저자	저술 관계 공동저자 관계	중심성 분석
2018	문성구, 김인재	논문, 저자	공동저자 관계	중심성 분석
2014	Waltman, Yan	논문	인용 관계	PageRank
2021	Weis, Jacobson	논문, 저자 학술지, 소속	인용 관계 저술 관계 공동저자 관계 소속 관계 게재 관계	PageRank
2021	Zhou 외	논문	인용 관계	중심성 분석 PageRank

한편 구축된 네트워크를 분석하는 방법 중에서 가장 널리 활용되는 것은 네트워크로 연결된 각 노드의 상대적 중요도를 나타내는 중심성 지수이다(이재윤, 2014). 이는 선행연구 대부분이 네트워크 분석방법으로 중심성 분석을 채택한 것을 통해 확인할 수 있다. 중심성 지수의 종류로는 연결 중심성 지수, 근접 중심성 지수, 매개 중심성 지수, PageRank 등이 활용된다.

네트워크에서 각 노드의 상대적 영향력을 나타내는 다양한 중심성 지수 중 논문의 영향력을 평가하기 위하여 최근 활용되는 방법이 PageRank이다. 대부분의 중심성 지수는 연결 관계의 빈도, 거리를 측정하여 중심성을 계산하지만, PageRank는 비율과 확률에 근거한다. PageRank(Brin & Page, 1998)는 각 웹페이지를 인용하는 웹페이지의 중요도를 산출하는 방법으로 ‘구글’이 사용자에게 검색결과 순위를 보여주는 방법에 활용되어 널리 알려져 왔다(이재윤, 2011).

Brin, Page(1998)가 제안하고 이후 연구자들에 의해서 수정된 웹페이지 A 의 PageRank 값은 다음과 같다.

$$PR(A) = \frac{1-d}{N} + d \left(\sum_{i=1}^N \frac{PR(T_i)}{C(T_i)} \right)$$

N : 전체 웹페이지의 수

T_i : 웹페이지 A 를 인용한 웹페이지(단, $1 \leq i \leq N$)

$PR(A)$: 웹페이지 A 의 PageRank

$C(T_i)$: 웹페이지 T_i 가 인용한 웹페이지의 수

d : damping factor 다른 노드로 이동할 확률

웹페이지 A 의 PageRank 값 $PR(A)$ 는 $PR(T_i)$ 의 값이 증가할수록 증가한다. 즉 PageRank 값이 높은 페이지 T_i 로부터 인용을 받으면 웹페이지 A 의 PageRank 값은 증가한다. 다시 말해 그 페이지를 인용하고 있는 다른 페이지의 PageRank 값이 높아야 한다. 웹페이지 A 의 PageRank 값 $PR(A)$ 는 $C(T_i)$ 의 값이 감소할수록 증가한다. 이것을 인용 네트워크에 적용하면 웹페이지 A 를 인용한 웹페이지 T_i 가 웹페이지 A 만 인용하고 다른 페이지는 인용하지 않아야 한다. 위의 2가지 사실을 종합하면 PageRank 값이 높은 웹페이지 T_i 가 A 만 인용하면 A 의 PageRank 값은 증가한다. 하지만 PageRank 값이 높은 웹페이지 T_i 가 A 뿐만 아니라 다른 웹페이지들도 인용했다면 PageRank 값은 감소한다. 이를 본 연구에 적용하면 영향력 있는 논문 T_i 가 참고문헌으로 선정된 논문 A 의 중요도를 높게 평가한다. 만약 영향력 있는 논문 T_i 가 다른 논문을 참고문헌으로 인용하지 않고 단 하나의 논문 A 만 인용했다면, 인용한 단 하나의 논문 A 는 그 영향력이 높게 계산될 것이다.

d 는 damping factor의 약자로 사용자가 검색결과를 만족하지 못하고 다른 페이지로 이동할 확률을 의미한다. Brin, Page(1998)는 $d=0.85$ 로 설정하였으며 이는 사용자가 85%의 확률로 검색결과에 만족하지 못하고 다른 페이지로 이동한다는 것을 의미한다. $0.85 \approx \frac{5}{6}$ 으로 근사할 수 있으며 사용자 입장에서는 대략 6번 정도 웹페이지를 탐색해야 1번 정도 원하는 결과를 얻을 수 있는 현실을 반영한 것으로 해석할 수 있다.

과학계량학에서는 학술 연구의 영향력을 PageRank 값으로 측정하려는 시도가 활발히 진행되고 있다. SCOPUS는 PageRank 값을 활용한 SCImago Journal Rankings(이하 SJR)를 학술지의 영향력 평가지표로 제공하고 있다. Web of Science도 PageRank 값을 활용한 Eigenfactor score를 학술지의 영향력을 평가지표로 서비스하고 있다(Mariani et al., 2016). 전 세계 연구자들이 가장 많이 참고하는 학술 DB에서 PageRank 값을 활용하여 연구의 영향력을 평가하고 있는 사례를 통해, PageRank 값은 연구의 영향력을 평가할 때, 피인용횟수 이상의 정보를 제공하는 것으로 인정되고 있음을 알 수 있다.

최근 과학계량학 분야의 연구(Bollen et al., 2006; Mariani et al., 2016; Weis & Jacobson, 2021)들은 학술지 뿐만 아니라 개별 논문의 영향력을 측정하고 예측하기 위해서 PageRank 값을 활용한다. 이는 해당 학문 분야를 네트워크로 구성하면 개별 논문을 둘러싸고 있는 환경에 대해서 현실과 유사하게 분석하고 결과를 도출할 수 있기 때문이다. Chen 외(2007)는 개별 논문을 평가하기 위해 PageRank 값을 활용하여 피인용횟수는 높지 않지만, 노벨상 수상에 기여한 중요한 논문을 판별하였다.

이번 절의 선행연구 분석 결과를 정리하면 다음과 같다. 과학계량학에서는 논문의 영향력을 피인용횟수로 정의하고 측정할 수 있지만, 여러 가지 한계로 인하여 새로운 접근법이 시도되고 있다. 새로운 방법으로는 해당 학문 분야를 네트워크로 구성하고, 네트워크의 중심성 지수를 활용하여 논문의 영향력을 측정하는 네트워크 분석 방법이 활용되고 있다. 네트워크의 중심성 지수는 여러 가지로 정의할 수 있지만, 현실을 잘 반영할 수 있는 PageRank 값으로 정의될 수 있음을 살펴보았다.

특히 인공지능으로 논문의 영향력을 예측하는 연구(Mariani et al., 2016; Weis & Jacobson, 2021)들은 해당 학문 분야의 네트워크를 구성하고, PageRank 값을 논문의 영향력으로 정의하고 있다. 나아가 선행연구들은 PageRank 값 상위 5% 이내 논문을 영향력 있는 논문으로 정의하였다. 이에 본 연구에서도 과학계량학의 관점에서 수학교육 논문의 영향력을 표현하였다. 먼저 수학교육 네트워크를 구성하고 논문 노드의 PageRank 값을 논문의 영향력으로 정의하였다. PageRank 값 상위 5% 이내의 논문을 수학교육의 영향력 높은 논문으로 정의하였다. 그리고 논문의 메타정보를 기계학습 모델에 입력하면, PageRank 값 상위 5% 이내 논문을 판별하는 모델을 설계하고자 하였다.

2. 연구방법 및 절차

가. 연구대상

이 연구의 대상은 수학교육 학술지에 게재된 개별 논문에 초점을 맞춘다. 수학교육의 연구는 연구를 수행하고 그 결과를 논문으로 작성하여 학술지에 게재하는 것으로 구성(Presmeg & Kilpatrick, 2019)되며 수학교육 연구의 결과는 논문을 통해서 확인할 수 있기 때문이다.

연구대상의 논문을 선정하기 위하여 수학교육 학술지를 평가한 연구(Nivens & Otten, 2017; Williams & Leatham, 2017)를 살펴보면, 학술지와 분석 기간을 선정한 후, 학술지에 게재된 논문을 분석대상으로 삼았다는 것을 알 수 있다. 본 연구도 중심성 지수 상위 5%를 기준으로 수학교육 논문을 분류하는 것이 목표이기에 연구대상인 수학교육 논문을 선정하기에 앞서 수학교육 논문이 게재되는 학술지를 선정하였다.

Nivens, Otten(2017)은 2010년부터 2013년까지 23개의 수학교육 학술지에 대하여 SCOPUS Rank를 기준으로 Q1-Q4까지 분류하였다. 한편 Williams, Leatham(2017)은 2가지 단계의 과정을 거쳐 학술지를 평가하였다. 첫째, 인용을 기반으로 한 SJR 점수로 분류하여 수학교육 연구에서 참고문헌으로 인용되는 비율이 98% 이상을 차지하고 있는 20개의 학술지를 분석대상으로 선정하였다. 20개의 학술지에 대하여 수학교육 전문가의 의견을 2개의 지역으로 분류하여 조사하여 [그림 II-1]과 같이 4개의 그룹으로 수학교육 학술지의 등급을 구분하였다.

Quality category	Region 1	Region 2
Very High	JRME, ESM	ESM, JRME
High	JMB, JMTE, ZDM, MTL, FLM	ZDM, JMB, FLM, MTL, JMTE
Medium High	RME, MERJ, IJSME, IJMEST, CJSMT, SSM, IJTME, JCMST, TME	IJSME, RME, MERJ, IJMEST, CJSMT, IJTME
Medium	IML, PRIMUS, TMA, TMME	JCMST, IML, SSM, TMA, TMME, TME, PRIMUS

[그림 II-1] 국외 수학교육 학술지의 순위(Williams & Leatham, 2017, p. 386)

이에 본 연구에서도 국외 학술지는 수학교육 연구에서 참고문헌으로 인용되는 비율이 98% 이상을 차지하고 있는 20개 학술지(Nivens & Otten, 2017)를 분석대상으로 삼았다.

한편 국내 수학교육 학술지의 영향력을 평가한 연구는 없었다. 이에 국내 수학교육 연구에 관한 메타분석 연구들을 살펴보면 국내 수학교육 관련 학술단체인 대한수학교육학회, 한국수학교육학회, 한국학교수학회에서 발행된 학술지를 중심으로 분석대상을 선정하고 있었다. 그 이유는 다른 학술단체의 학술지도 있었지만, 앞의 세 학술단체는 전국 규모의 학술대회를 개최하는 학회로서 국내의 수학교육 분야에서 권위를 인정(박선영, 김원경, 2011)받고 있기 때문이었다. 본 연구는 위의 학술단체의 학술지를 중심으로 폭넓은 수학교육 연구를 메타분석하기 위하여 KCI 등재 후보지까지 포함하여 총 9종의 수학교육 학술지에 게재된 논문을 분석대상을 선정하였으며 그 결과는 <표 II-2>와 같다. 분석대상의 학술지는 ESM, JRME를 포함한 국외 20개, ‘수학교육’을 포함한 국내 학술지 9개, 총 29개로 구성되었으며, 학술지의 지역을 정리한 결과 미국 10개, 유럽 10개, 한국 9개로 균형 있게 선정되었다.

<표 II-2> 분석대상 학술지

학술지명 (약자)	ISSN	창간연도	지역	국내/국외
수학교육(The Mathematical Education)	12251380	1963	Republic of Korea	국내
한국수학사학회지(Journal for History of Mathematics)	1226931X	1984	Republic of Korea	국내
수학교육 논문집(Communications of Mathematical Education)	12266663	1990	Republic of Korea	국내
수학교육학연구(The Journal of Educational Research in Mathematics (JERM))	22887733	1991	Republic of Korea	국내
초등수학교육(Education of Primary School Mathematics)	12266914	1997	Republic of Korea	국내
수학교육 연구(Research in Mathematical Education)	12266191	1997	Republic of Korea	국내
한국초등수학교육학회지(Journal of Elementary Mathematics Education in Korea)	12293229	1997	Republic of Korea	국내
한국학교수학회(Journal of the Korean School Mathematics)	12290890	1998	Republic of Korea	국내
학교수학School Mathematics (SM)	12294322	1999	Republic of Korea	국내
Educational Studies in Mathematics (ESM)	00131954, 15730816	1968	Netherlands	국외
International Journal of Mathematical Education in Science and Technology (IJMEST)	0020739X, 14645211	1970	United Kingdom	국외
School Science and Mathematics (SSM)	00366803, 19498594	1973	United States	국외
Journal of Computers in Mathematics and Science Teaching (JCMST)	07319258	1981	United States	국외
Teaching Mathematics and Its Applications (TMA)	2683679, 14716976	1982	United Kingdom	국외
Mathematics Education Research Journal(MERJ)	10332170, 2211050X	1989	Netherlands	국외
The Mathematics Educator (TME)	10629017	1990	United States	국외
PRIMUS (Problems, Resources and Issues in Mathematics Undergraduate Studies)	10511970, 19354053	1991	United Kingdom	국외
Journal of Mathematical Behavior (JMB)	07323123	1994	United States	국외
Journal for Research in Mathematics Education (JRME)	00218251	1996	United States	국외
Zentralblatt für Didaktik der Mathematik (ZDM)	18639690, 18639704	1997	Germany	국외
Research in Mathematics Education (RME)	14794802, 17540178	1998	United States	국외
Canadian Journal of Science, Mathematics and Technology Education (CJSMTE)	14926156, 19424051	2001	United States	국외
International Journal of Science and Mathematics Education (IJSME)	15710068	2003	Netherlands	국외
Journal of Mathematics Teacher Education (JMTE)	13864416	2005	Netherlands	국외
Investigations in Mathematics Learning (IML)	19477503, 24727466	2008	United States	국외
Mathematical Thinking and Learning (MTL)	10986065, 15327833	2009	United States	국외
For the Learning of Mathematics (FLM)	02280671	2011	Canada	국외
International Journal for Technology in Mathematics Education (IJTME)	17442710, 20452519	2012	United Kingdom	국외
The Montana Mathematics Enthusiast(TMME)	15513440	2013	United States	국외

나. 연구대상 논문의 메타정보 수집

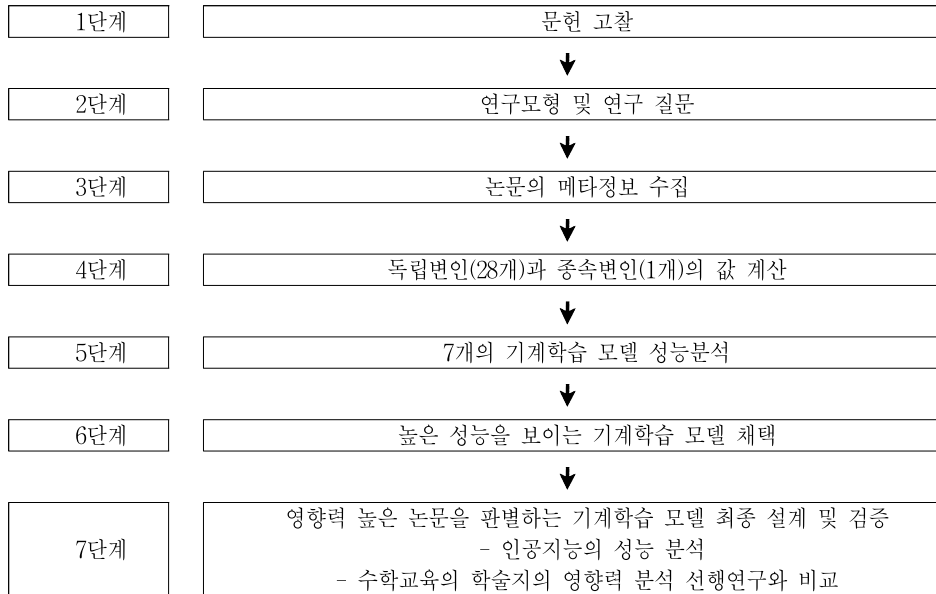
연구대상에 해당하는 논문들의 메타정보는 PubMed, Crossref, Microsoft Academic, CORE 및 PubMed Central의 통합되는 LENS.ORG API에서 검색된 데이터를 사용하여 구성하였다. LENS.ORG는 학술 DB로 논문의 제목, 저자, 학술지에 대한 정보뿐 아니라 논문에 명시된 참고문헌을 추출하여 제공하고 있다. LENS.ORG에서 검색된 데이터에는 <표 II-2>에서 정리된 29개의 수학교육 관련 학술지의 논문이다. 검색일(2022년 8월)기준으로, 27,012편의 논문이 검색되었다. 방정숙 외(2019)에 따르면 1963년부터 1980년까지 수학교육 연구는 체계적인 논문의 형식을 갖추지 못한 경우도 있었고, ‘수학교육’에 대한 연구보다는 ‘수학’에 대한 연구의 수가 더 많았다. 이에 수학교육 논문 중 1980년에서 2021년 사이에 발표된 논문을 분석대상으로 선정하였다.

처음 API를 활용하여 논문의 메타데이터를 추출할 때, <표 II-2>학술지의 ISSN을 활용한 결과 약 8천 건 정도만 추출되었다. 그 이유는 LENS.ORG의 데이터 중에서 ISSN이 누락된 것이 많기 때문이었다. 이에 학술지명을 기준으로 2022년 8월에 다시 추출하였다. LENS.ORG는 논문의 판별기준으로 LensID를 활용하였기에 본 연구에서도 LensID를 활용하였다. 한편 논문의 저자 정보에 대한 판별기준은 Microsoft Academic ID(이하 MagID)와 orcid를 활용하고 있었다. LENS.ORG가 제공하는 ORCID는 70% 이상 누락되어 있었기 때문에, 10% 정도 누락률을 보이는 Microsoft Academic ID를 활용하였다. 데이터의 전처리 과정에서 저자의 판별 정보가 누락되거나 참고문헌이 29개 학술지가 아닌 논문들은 제외하였다. 네트워크의 PageRank 값으로 중심성 지수를 계산하기 위하여 29개 학술지의 논문을 1편 이상 참고문헌으로 선택한 논문만 남기는 전처리를 하였다. 전처리는 NEO4J와 파이썬을 이용하여 진행되었으며 저자 정보가 누락되었거나, 29개 이외의 학술지에서 참고문헌을 선택한 논문을 제외하여 데이터를 전처리한 결과 27,012편의 논문 중 24,575편의 논문에 대한 메타데이터 정보를 정리할 수 있었다.

API를 이용하여 총 9개의 항목을 자동으로 수집하는 코드를 작성하고 수집하였다. 논문의 메타정보와 관련하여 논문의 판별기준(LensID), 논문의 제목(title), 출판 연도(year_published), 참고문헌(references)의 판별 정보(LensID) 총 5개의 항목을 수집하였다. 저자의 메타정보와 관련하여 저자명(authors) 저자의 소속기관명(authors.affiliations.name) 저자판별 정보(authors.ids) 총 3개의 항목을 수집하였다. 학술지와 관련된 메타정보는 학술지명(source.title)과 학술지가 속한 국가명(source.country) 2개의 항목을 수집하였다. 연구대상의 출판 연도는 1980년도부터 2021년까지 42년간이었다. 메타정보의 수집은 파이썬으로 코드를 작성하고 LENS.ORG의 API를 이용하여 진행하였다. 수집 후에는 학술지별 폴더에 5개의 파일을 생성하여 저장하였다.

다. 연구 방법 및 절차

영향력 있는 논문을 판별하는 기계학습 모델을 설계하는 절차를 서술하면 [그림 II-2]와 같다. 먼저 ‘논문의 영향력을 평가’, ‘기계학습 분류모델’에 대한 문헌을 고찰하여 논문의 영향력에 영향을 미치는 독립변인을 도출하였다. 다음 단계로 독립변인에 관한 데이터 수집을 위하여 LENS.ORG에서 논문의 메타정보를 수집하였다. 수집한 메타정보로부터 독립변인과 종속변인의 정의에 해당하는 값을 파이썬을 활용하여 산출하며 데이터를 전처리하였다. 7종류의 기계학습 모델에 독립변인과 종속변인을 입력하여 성능을 비교 분석하였으며, PYTHON을 활용하였다. 기계학습으로 분류모델 설계한 선행연구를 고찰한 결과에 근거하여 7개의 기계학습 분류 모델(로지스틱 회귀, KNN, 나이브 베이즈, 의사결정트리, 랜덤포레스트, lightGBM, LinearSVC)을 선정하였다. 이 7개 모델 중에서 최적의 모델을 찾기 위하여 각각 모델을 설계하고 그 성능을 비교하였다. 성능을 비교하기 위한 지표로 정확도(accuracy), 오 분류율(error rate), 정밀도(precision), 재현율(recall), 특이도(specificity), FP Rate(False Positive Rate)를 활용하여 분석하였다. 그중에서 가장 좋은 성능을 보이는 모델을 채택하였다. 설계한 기계학습 모델을 검증하기 위하여 분류 성능을 비교하고 수학교육 학술지를 평가했던 선행연구 결과와 비교하였다.



[그림 II-2] 기계학습 모델 설계 절차

라. 기계학습과 인공지능 분류모델 성능측정

기계학습은 인공지능의 한 분야로 데이터를 학습하여 새로운 결과를 예측하거나 데이터를 분류한다. 학습데이터(training data)에 종속변인(label)이 있는 경우를 지도학습이라고 하고, 종속변인이 없는 경우를 비지도학습이라고 한다.

분류는 대상을 정해진 범주로 나누는 것을 의미한다. 지도학습으로 데이터를 분류하는 과정은 먼저 종속변인이 포함된 데이터를 학습하여 모델을 생성한다. 입력된 데이터의 종속변인은 범주화되어 있으므로 기계학습 모델이 출력하는 결과값도 학습된 데이터의 종속변인 범주 중 하나이다. 생성된 모델에 새로운 데이터가 입력되면, 어느 범주의 종속변인에 속하는지 찾는 방식으로 작동한다.

종속변인이 있는 데이터를 학습하여 분류하는 인공지능 모델의 성능을 평가하는 지표로 정 분류율, 오 분류율을 활용할 수 있다. 정 분류율은 Positive(이하 P) 또는 Native(이하 N)라는 범주로 분류된 종속변인에 대하여 P를 P로 분류하고, N을 N으로 분류하는 비율을 의미한다. 마찬가지로 오 분류율은 P를 N으로, N을 P로 분류한 비율을 뜻한다. 분류모델은 <표 II-3>과 같은 혼동행렬을 이용하여 표현할 수 있다.

<표 II-3> 혼동행렬

	실제값(P)	실제값(N)
예측값(P)	True Positive(이하 TP)	False Positive(이하 FP)
예측값(N)	False Negative(이하 FN)	True Negative(이하 TN)

머신러닝이 P를 예측하면 Positive, N을 예측하면 Negative이라 하고, 머신러닝의 예측결과가 맞다면 True, 틀리다면 False로 표기한다. 이를 바탕으로 모델의 성능을 평가하기 위한 성능지표는 <표 II-4>와 같이 정리할 수 있다.

<표 II-4> 분류모델 성능지표

항목	계산식	설명
정확도 (accuracy)	$\frac{TP + TN}{TP + FP + FN + TN}$	전체 결과 중 맞게 분류한 비율
정밀도 (precision)	$\frac{TP}{TP + FP}$	P로 예측한 것 중 실제 P인 비율
재현율 (sensitivity, recall)	$\frac{TP}{TP + FN}$	실제 P를 P로 예측한 비율
FP Rate (False Alarm Rate)	$\frac{FP}{FP + TN}$	실제 N이 아닌 것을 N으로 예측한 비율

정확도(accuracy)는 실제값과 예측값이 일치하는 비율을 뜻하며 본 연구에서는 수학교육 네트워크 중심성 지수 상위 5% 이내이면 Positive, 그 외에는 Negative로 설정하였다. 이때, 실제 Positive와 Negative는 수학교육 네트워크에서 중심성 지수 상위 5% 이내인지 아닌지로 결정할 수 있으며, 예측 Positive와 Negative는 머신러닝 모델이 예측한 분류 결과로 한다.

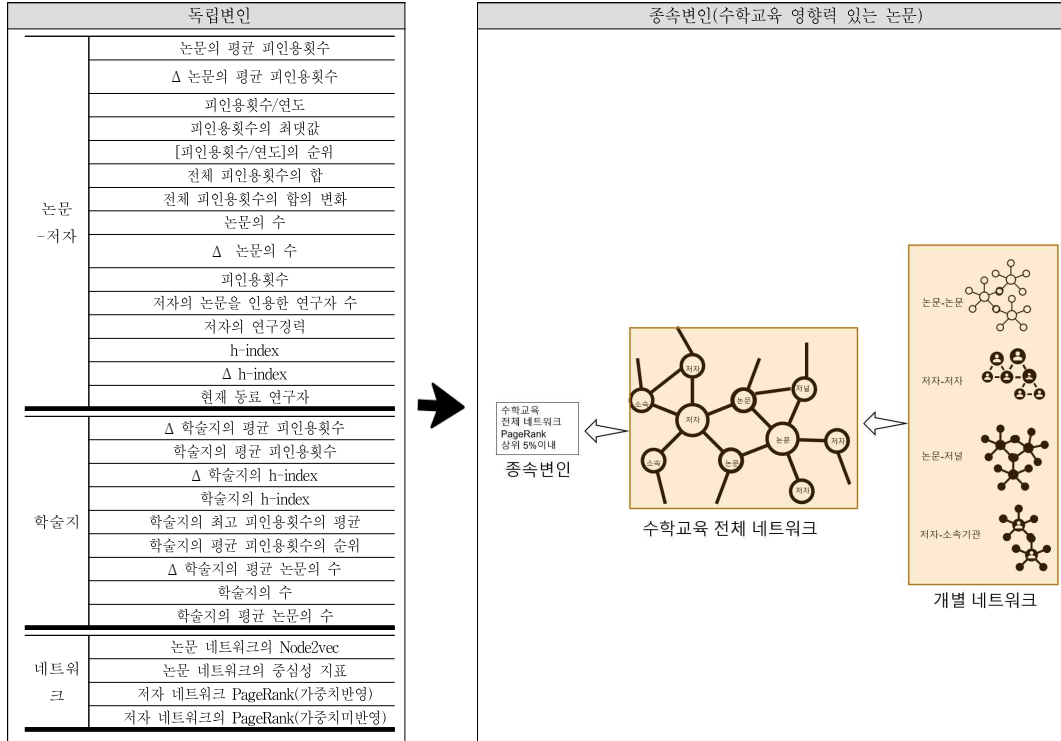
정밀도는 N을 P로 잘못 판단하는 경우 큰 문제가 발생하는 영역에서 활용한다. 예를 들어 스팸메일을 P, 일반메일을 N으로 정의하면, 일반메일을 스팸으로 잘못 판단하면 중요한 메일을 확인하지 못할 수 있기에 스팸메일 분류 머신러닝 모델은 정밀도를 확인하여 성능을 평가한다. 본 연구에서 정밀도는 실제 네트워크 중심성 지수 상위 5% 이내가 아니지만, 인공지능 모델은 상위 5% 이내라고 분류한 논문들의 비율을 뜻한다.

재현율은 P를 N으로 잘못 판단하는 경우 큰 문제가 발생하는 영역에서 활용한다. 예를 들어 의료분야에서 실제 질병을 P, 건강한 상태를 N으로 설정하면 실제 질병이 있지만, 질병이 없다고 판단하면 생명과도 직결된 문제가 발생하기 때문에 재현율을 의미 있게 점검한다. 본 연구에서는 실제 네트워크 중심성 지수 5% 이내이지만, 인공지능 모델은 네트워크 중심성 지수 5% 이내가 아니라고 분류한 논문들을 뜻한다.

III. 연구 결과 및 논의

1. 수학교육 연구의 영향력 판별 모형

본 연구는 학술 연구를 네트워크로 구성하여 인공지능으로 각 논문의 영향력을 예측한 Weis, Jacobsen(2021)의 모형을 기반으로 수학교육 연구의 특성을 고려하여 아래의 [그림 III-1]과 같이 수정한 모형을 설정하였다.



[그림 III-1] 기계학습 모델 설계모형

과학계량학에서 연구의 영향력 평가는 피인용횟수 기반의 전통적인 모델에서 인공지능을 활용하여 평가하는 방법으로 발전하고 있다. 이 중에서 Weis, Jacobson(2021)의 모형을 활용하여 해당 학문 분야를 네트워크로 구성하면 개별 논문을 둘러싸고 있는 환경에 대해서 현실과 유사하게 분석하고 결과를 도출할 수 있다. 이에 수학교육 분야에서 영향력 있는 논문을 판별하기 위하여 Weis, Jacobson(2021)의 모형을 활용하였다.

수학교육의 영향력 높은 논문을 판별하기 위한 기계학습 모델은 지도학습 기반으로 설계하였기에 입력데이터를 독립변인과 종속변인으로 구성하였다. 종속변인은 각 논문에 대하여 수학교육 영향력 높은 논문으로 PageRank 값 상위 5% 이내이면 1, 아니면 0으로 설정하였다. 한편 독립변인은 각 논문의 메타정보로부터 계산할 수 있는 값으로, Mariani 외(2016)과 Weis, Jacobson(2021)의 연구에서 영향력(종속변인)을 예측하기 위하여 사용했던 독립변인(28개)으로 구성하였다. 측정 변인의 값 산출은 Python 3.9.7과 NEO4J 4.4.5 환경에서 이루어졌다. 본 연구의 목적은 수학교육의 영향력 높은 논문을 판별하기 위하여 기계학습 모델을 설계하는 것이다. 구체적으로는 논문이 출판된 지 5년 동안의 독립변인을 토대로 출판된 지 5년이 되는 시점에서 네트워크 중심성 지수가 상위 5% 이내인 논문을 판별하는 것이다. 이를 위하여 학술 연구의 영향력 평가에 관한 선행연구, 인공지능으로 논문의 영향력을 예측하는 문헌을 고찰하여 네트워크 중심성 지수에 영향을 미치는 28개 독립변인을 도출하였다. 독립변인으로 구성된 메타정보는 인공지능으로 영향력 있는 학술연구를 예측한 연구(Mariani et al., 2016; Weis & Jacobson, 2021)를 근거로 3가지 범주(논문-저자, 학술지, 네트워크)로 구분할 수 있다.

각 변인에 해당하는 값 중에는 h-index와 같이 학술 DB에서 바로 수집할 수 있는 항목도 있었지만, 학술 DB

에서 제공하지 않는 항목이 대부분이었다. 또한, 본 연구의 연구대상인 수학교육의 29개 학술지가 아닌 다른 학술 분야의 학술지에 게재된 논문과 인용도 고려된 값이다. 이에 본 연구의 분석대상인 수학교육 학술 연구의 네트워크 내의 연결정보를 통해서 해당하는 값을 산출하기 위하여, <표 III-1>과 같이 API를 이용하여 수집된 항목 “논문 식별 정보(LensID), 출판 연도, 저널명, 저자식별정보(MagID), 참고문헌 식별 정보(LensID), 저널명”을 기준으로 수식을 만들고 파이썬을 이용하여 계산하였다.

<표 III-1> API수집 데이터 양식 예시

논문 식별정보 (LensID)	출판 연도	저자식별정보 (MagID)	참고문헌 식별정보 (LensID)	저널명
000-404-304-872-870	2006	M2583937674, M2091503171, M2675631524	005-500-506-031-537, 015-816-976-088-902	A
019-197-618-091-576	2014	M2583937674	000-404-304-872-870	A
...
017-061-197-276-031	2020	M2028622666	019-197-618-091-576, 022-180-533-751-552, ...	B

예를 들어 ‘저자의 연구 경력’은 저자의 첫 번째 논문이 출간된 연도를 기준으로 계산되었으며, ‘논문의 수’는 저자식별정보를 기준으로 논문 식별 정보(LensID)의 수를 합하여 계산하였고, ‘현재 동료 연구자’는 저자가 출판한 논문 식별 정보(LensID)를 기준으로 공동 저자의 식별 정보(MagID)의 수를 계산하였다. ‘피인용횟수의 최댓값’은 ‘저자가 학술지에 게재된 논문이 받은 피인용횟수의 최댓값’을 의미(Weis, Jacobson, 2021)하였기에 저자의 논문이 받은 피인용횟수의 값을 먼저 확인하고, 이 중 최댓값을 구하였다. 만약 공동저자가 있는 경우는 각 저자의 피인용횟수 최댓값을 평균하여 계산하였으며, 독립변인의 계산 과정은 [그림 III-2]처럼 파이썬을 활용하였다. 다른 독립 변인들도 Weis, Jacobson(2021)에서 제시한 정의에 근거하여 파이썬을 이용하여 계산하였다.

4. Maximum citations

```

1 #4.1 저자의 연도별 인용수 최댓값
2 magid_max_cit={}
3 for mg in tqdm(magid_lensid_year):
4     m_max_c=np.zeros(42)
5     for l_y in magid_lensid_year[mg]:
6
7         if lensid_year_sum_cit.get('%s' %l_y) is None:
8             m_max_c += 0
9         else:
10            for i in range(42):
11                if m_max_c[i]<lensid_year_sum_cit.get('%s' %l_y)[i]:
12                    m_max_c[i]=lensid_year_sum_cit.get('%s' %l_y)[i]
13            magid_max_cit['%s'%mg]=m_max_c
14
15 max_len_max_cit={}
16 max_len_max_cit=lensid_metric(max_len_max_cit, magid_max_cit)
17 max_len_max_cit_Syear={}
18 max_len_max_cit_Syear=lensid_metric_Syear(max_len_max_cit_Syear, max_len_max_cit)
19
20 filePath = './data/metric4_len_max_cit_Syear.txt'
21 with open(filePath, 'wb') as f:
22     pickle.dump(max_len_max_cit_Syear, f)

```

[그림 III-2] 독립변인의 값을 계산하기 위한 파이썬 코드의 일부

본 연구에서는 총 24,575편의 논문에 대하여 논문마다 28개의 독립변인에 해당하는 값을 논문이 출판된 이후 5년 동안 매년 계산하여 최종 140개의 독립변인과 1개의 종속변인으로 입력데이터를 구성하였다. 기계학습 모델에 입력할 데이터는 아래의 [그림 III-3]과 같이 표 형식의 데이터로 행은 연구대상 논문의 총 개수인 24,575개였으며, 열은 독립변인 140개와 종속변인 1개로 구성되어 총 141개였다. 한편 2020년 이후 게재된 논문의 경우, 출판 2년후의 정보부터는 수집할 수 없으므로 해당 값을 0으로 처리하였다.

논문 판별 정보	출판 당해연도 독립변인					...	출판 4년 후 독립변인					중속변인 5% 여부
	1	2	3	...	28		...	1	2	3	...	
1												
2												
3												
...												
24575												

[그림 III-3] 입력데이터 형식

2. 수학교육 연구의 영향력 판별 기계학습 모델 성능 비교결과

수학교육 연구의 영향력 판별 기계학습 모형에 적합한 기계학습 모델을 선택하기 위하여 같은 전처리 방법으로 생성한 입력데이터를 활용하였다. 구체적으로는 수학교육 네트워크로부터 계산된 140개의 독립변인과 1개의 종속변인으로 입력데이터를 생성하였다. 이를 7개의 기계학습 분류모델에 각각 입력하여 성능을 분석한 결과는 <표 III-2>와 같다.

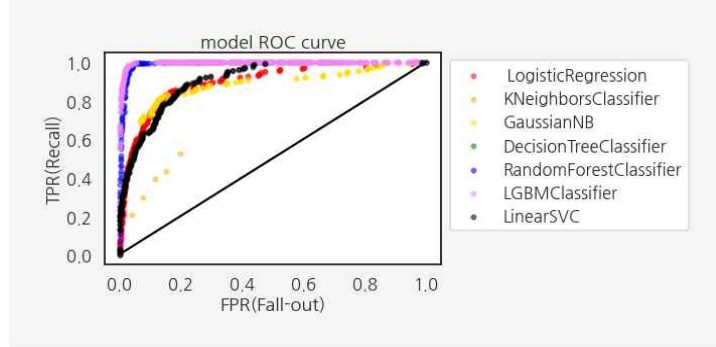
기계학습 모델이 분류한 결과와 실제값이 일치하는 여부를 확인할 수 있는 정확도(Accuracy)는 모든 분류모델이 0.83 이상이었지만, 랜덤포레스트 분류모델이 0.980으로 가장 높았다. 정밀도(Precision)는 인공지능이 상위 5% 이내라고 예측한 논문 중 실제로 상위 5% 이내에 속하는 논문의 비율에 해당하는 값으로 기계학습 모델마다 0.111부터 0.770까지 큰 차이를 보였으며, lightGBM 모델은 0.770으로 가장 큰 값을 보였으며, 랜덤포레스트 분류모델이 0.726으로 두 번째로 높은 값을 보였다. 재현율(Recall)은 실제 상위 5% 이내의 논문을 인공지능이 상위 5% 이내로 예측하는지 확인할 수 있는 값으로 기계학습 모델마다 0.309부터 0.883까지 큰 차이를 보였으며, 가장 높은 성능을 보인 분류모델은 서포트 벡터 머신이었으며, 랜덤포레스트 모델은 0.860으로 두 번째로 큰 값을 보였다.

<표 III-2> 인공지능 분류모델 성능분석(평균값)

인공지능 분류모델	Accuracy	Precision	Recall (TPR)	Fallout (FPR)	AUC
로지스틱 회귀	0.839	0.177	0.729	0.156	0.787
KNN	0.861	0.111	0.309	0.113	0.598
나이브 베이즈	0.851	0.202	0.812	0.148	0.823
의사결정트리	0.966	0.593	0.702	0.021	0.839
랜덤포레스트	0.980	0.726	0.860	0.014	0.922
lightGBM	0.976	0.770	0.797	0.014	0.892
서포트 벡터 머신	0.867	0.258	0.883	0.136	0.792

한편 ROC 곡선은 True Positive(TP) 비율에 대한 False Positive(FP) 비율을 평면에 표현한 것이며 분류모델 별로 ROC 곡선은 [그림 III-4]와 같다. 그림의 중앙의 검은선인 $y=x$ 와 각 모델의 ROC 곡선이 이루는 넓이가 넓을수록 성능이 우수한 것으로 볼 수 있다. 파란색 점으로 표현된 랜덤포레스트 분류모델과 분홍색 점으로 표현된 lightGBM 분류모델이 다른 분류모델에 비해서 성능이 우수함을 확인할 수 있다.

또한, ROC 곡선 아래의 넓이는 AUC라 하며, AUC 값이 클수록 높은 성능을 보인다. <표 III-2>와 같이 각 인공지능 분류모델별 AUC 값은 0.598부터 0.922까지 분포를 보였으며, 랜덤포레스트 모델이 0.922로 가장 높은 성능을 보였다.



[그림 III-4] 분류모델별 ROC 곡선

기계학습 분류모델의 성능 분석 결과, 랜덤포레스트 분류모델이 대부분 성능지표의 값이 가장 컸다. 재현을 항목은 7개의 분류모델 중 서포트 벡터 머신 분류모델이 가장 큰 값을 보였지만, 정밀도의 값이 0.258로 성능지표에 따라 값의 편차가 크게 나타났다. 모든 성능지표에서 높은 값을 보이면서도 성능지표에 따른 편차가 크지 않은 랜덤포레스트를 분류모델로 채택하였다.

3. 수학교육 연구의 영향력 판별 기계학습 모델 검증

본 연구에서 중심성 지수 상위 5% 이내 논문을 분류하는 기계학습 모델에 수집했던 24,575편의 논문 메타정보를 모두 입력하여 그 결과를 확인하였다. 랜덤포레스트 모델은 분석대상의 논문 24,575편 중 1,017편의 논문이 중심성 지수 5% 이내일 것으로 예상하였다. 기계학습 모델이 산출한 결과의 타당성을 분석하기 위하여 상위 5% 이내로 예측된 논문의 비율을 기준으로 상위 15개의 학술지를 <표 III-3>과 같이 정리하였다.

<표 III-3> 중심성 지수 5%이내 분류 결과(상위 15개 학술지)

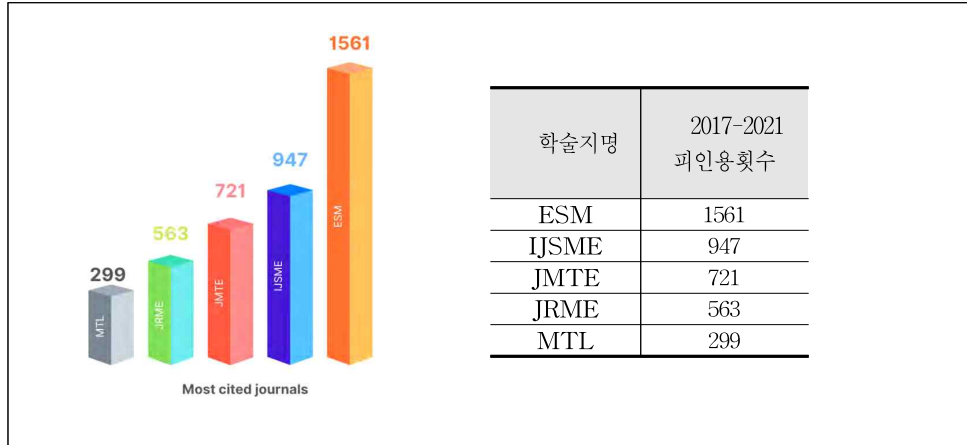
학술지명	국내/국외	중심성 지수 5% 이내 논문의 비율	전체 논문의 비율
JRME	국외	25.66	8.23
ESM	국외	22.12	8.82
ZDM	국외	8.46	5.49
JMTE	국외	5.8	2.73
JMB	국외	5.51	4.13
IJSME	국외	5.51	5.29
PRIMUS	국외	4.92	6.22
MTL	국외	4.23	1.77
SSM	국외	4.03	12.46
FLM	국외	3.24	2.6
IJMEST	국외	2.75	17.24
TMME	국외	1.28	2.86
RME	국외	1.08	2.04
MERJ	국외	0.98	3.18
수학교육 논문집	국내	0.88	0.54

JRME, ESM이 25.66%, 22.12%로 가장 높은 비율을 보인다. ESM과 JRME은 수학교육 분야의 연구자들에게 주요 학술지로 인정(Inglis & Foster, 2018)받고 있는 현실과 연구결과는 부합한다. 또한, Williams, Leatham(2017)은 미국 지역과 미국이 아닌 지역의 수학교육 연구자들에게 수학교육 학술지의 영향력을 질적으로 평가하게 하였고, 4단계로 분류하였다. Williams, Leatham(2017)의 학술지의 영향력 평가결과를 인공지능 분류모델의 결과와 비교하기 위하여 상위 5% 이내에 속하는 논문이 10편 이상이 있는 학술지를 <표 III-4>와 같이 정리하였다. Williams, Leatham(2017)의 연구결과에 따르면 JRME과 ESM은 두 지역 연구자들에게 모두 'very high' 평가를 받았다. ZDM, JMB, For the Learning of Mathematics(이하 FLM), MTL, JMTE는 두 지역의 연구자들이 공통적으로 'high'로 평가하였다. Williams, Leatham(2017)의 연구결과는 본 연구에서 설계한 인공지능 모델의 분류 결과와 상당 부분 일치한다. 인공지능 분류모델이 도출한 영향력 있는 논문의 학술지별 비율은 JRME 25.66%, ESM 22.12%, ZDM 8.46%, JMTE 5.8%, JMB 5.51%, Problems, Resources and Issues in Mathematics Undergraduate Studies(이하 PRIMUS) 4.92%, MTL 4.23%, School Science and Mathematics(이하 SSM) 4.03%, FLM 3.24% 순이었다. 다만, Williams, Leatham(2017)의 연구결과에서 수학교육 연구자들에게 'medium', 'medium high' 평가를 받은 학술지 중 PRIMUS, SSM, IJSME에 대하여 인공지능은 그 영향력을 높게 평가하고 있는 것이 확인되었다. 이는 학술지의 명성은 높지 않지만, 역량이 우수한 개별 연구자의 성과물이 해당 학술지에 게재되고 있음을 추론할 수 있다.

<표 III-4> 본 연구의 결과와 Williams, Leatham(2017)의 결과와 비교

학술지명	본 연구의 결과 영향력 있는 논문의 비율	Williams, Leatham(2017) 학술지 영향력 질적 평가 등급	
		Region1	Region2
JRME	25.66	very high	very high
ESM	22.12	very high	very high
ZDM	8.46	high	high
JMTE	5.8	high	high
JMB	5.51	high	high
IJSME	5.51	medium high	medium high
PRIMUS	4.92	medium	medium
MTL	4.23	high	high
SSM	4.03	medium high	medium
FLM	3.24	high	high
IJMEST	2.75	medium high	medium high
TMME	1.28	medium	medium
RME	1.08	medium high	medium high
MERJ	0.98	medium high	medium high

한편 Yig(2022)는 2017년부터 2021년까지 수학교육 학술지의 인용횟수를 조사하였고 [그림 III-5]와 같이 정리하였다. Yig(2022)의 연구에 따르면 International Journal of Science and Mathematics Education(이하 IJSME)이 JRME보다 더 많은 인용을 받고 있었으며, ESM 다음으로 많이 인용되는 학술지로 조사되었다. 이러한 결과는 과학계량학의 관점에서 피인용횟수만으로 학술 연구의 영향력을 평가할 때 발생하는 문제점을 잘 보여주는 사례이다.



[그림 III-5] 수학교육 학술지별 피인용횟수 순위(Yig, 2022, p. 143)

Yig(2022)의 연구결과는 두 학술지의 목적과 범위(aim and scope)를 근거로 분석할 수 있다. 학술지 IJSME의 목적 및 범위를 확인하면 “The objective of this journal is to publish original, fully peer-reviewed articles on a variety of topics and research methods in both **science and mathematics education.**”으로 과학교육과 수학교육 모두에서 다양한 주제를 다루고 있다. 따라서 수학교육만을 다루는 JRME와 피인용횟수의 차이가 있는 것은 당연하다. 또한, 피인용횟수는 IJSME이 JRME보다 많지만, 여전히 수학교육 연구자들은 JRME의 영향력을 IJSME보다 높게 인식하고 있다. 이는 두 학술지의 JCR에서 제공한 h-index를 확인해도 마찬가지이다. JRME의 h-index는 82이고, IJSME의 h-index는 45이다. h-index 값을 해석하면 JRME는 피인용횟수가 82 이상인 논문이 82편 이상 게재되었으며, IJSME는 피인용횟수가 45 이상인 논문이 45편 게재된 것을 의미한다.

학술지에 많은 논문이 게재되고, 그 논문이 각각 1 혹은 2의 피인용횟수를 받아도 피인용횟수의 누적값은 증가한다. 하지만, 영향력 있는 학술지는 다른 연구자들이 관심을 보이며 인용을 많이 받는 논문이 게재되어야 한다. 이처럼 피인용횟수의 순위로 IJSME와 JRME를 평가할 수 없음을 다시 한번 확인할 수 있다.

이에 비해 본 연구의 기계학습 모델이 예측한 상위 5% 이내의 논문 중 JRME에 게재된 논문은 25.66%, 반면 IJSME에 게재된 논문은 5.51%로 약 20% 이상의 차이를 보인다. 이를 통해 본 연구에서 설계한 기계학습 모델은 피인용횟수만으로 평가할 수 없는 연구의 영향력을 잘 판별하고 있음을 확인할 수 있다.

본 연구에서는 중심성 지수 5% 이내를 분류하는 기계학습 모델의 결과를 해석하기 위하여 중심성 지수 10% 이내, 중심성 지수 20% 이내 분류모델을 추가 생성하였다. PageRank 값을 기준으로 상위 10% 혹은 상위 20%에 해당하는 점수를 산출하고, 이 기준에 맞게 입력데이터의 종속변수를 다시 범주화하여 구성하였다. 이렇게 수정된 입력데이터를 랜덤포레스트 모델이 학습하여 중심성 지수 상위 10% 이내, 중심성 지수 상위 20% 이내를 분류하는 모델을 각각 생성하였다. 그리고 생성한 분류모델을 활용하여 중심성 지수 상위 10% 이내, 중심성 지수 상위 20% 이내의 수학교육 논문을 판별하고 정리한 결과는 <표 III-5>와 같다.

중심성 지수 상위 5%, 10%, 20% 이내로 분류된 논문이 게재된 학술지는 중심성 지수의 %에 상관없이 전체적으로 비슷한 양상을 보였다. 중심성 지수를 상위 5% 이내 → 상위 10% 이내 → 상위 20% 이내로 완화하였지만, 중심성 지수가 높은 논문들이 게재된 학술지는 ESM과 JRME였다. 다만 분류 기준을 완화할수록 두 학술지의 점유율은 47.78%, 39.47%, 31.64%로 감소하고 있었다.

<표 III-5> 중심성 지수 5%, 10%, 20%이내 분류 결과

저널 명칭	국내/국외	전체 데이터	상위 5% 이내	상위 10% 이내	상위 20% 이내
JRME	국외	8.23%	25.66%	19.29%	14.4%
ESM	국외	8.82%	22.12%	20.18%	17.24%
ZDM	국외	5.49%	8.46%	8.34%	8.22%
JMTE	국외	2.73%	5.8%	5.33%	4.53%
JMB	국외	4.13%	5.51%	5.48%	5.76%
IJSME	국외	5.29%	5.51%	7.01%	7.54%
PRIMUS	국외	6.22%	4.92%	6.17%	6.18%
MTL	국외	1.77%	4.23%	4.49%	3.52%
SSM	국외	12.46%	4.03%	4.34%	5.33%
FLM	국외	2.6%	3.24%	3.6%	3.55%
IJMEST	국외	17.24%	2.75%	4.74%	6.41%
TMME	국외	2.86%	1.28%	1.58%	2.09%
RME	국외	2.04%	1.08%	0.84%	1.22%
MERJ	국외	3.18%	0.98%	1.33%	2.37%
수학교육 논문집	국내	0.54%	0.88%	1.13%	1.39%

마찬가지로 중심성 지수의 기준을 완화하여도 중심성 지수 상위 논문들이 게재된 학술지는 8개 학술지(ZDM, JMB, JMTE, IJSME, PRIMUS, MTL, SSM, FLM)가 총 40%에 가까운 점유율을 보였다. 한편 International Journal of Mathematical Education in Science and Technology(이하 IJMEST)는 중심성 지수가 완화될수록 2.75% → 4.74% → 6.41%로 지속적으로 증가하는 모습을 보였다. 학술지 IJMEST의 목적과 범위와 관련된 내용 중 “Mathematical models arising from real situations, the use of computers, new teaching aids and techniques also form an important feature.”를 통해서 알 수 있듯이, IJMEST는 컴퓨터, 새로운 도구, 기술에 주목한다. IJMEST의 결과를 통해서 수학교육 연구의 흐름과 양상을 발견할 수 있다. 수학교육 연구 분야에서 과학기술을 활용한 수학교육 연구들은 상위 5%로 분류되는 비율은 낮지만, 과학기술의 발전과 급변하는 교육 환경으로 인하여 상위 20%로 분류될 가능성은 높은 것으로 짐작할 수 있다.

이상의 연구결과를 정리하면 영향력 있는 논문을 판별하는 모델에 데이터를 입력하여 가장 적합한 모델을 확인하였다. 모델에 대한 검증은 성능을 비교분석하는 인공지능의 방법과 실제 수학교육 연구자들이 수학교육 학술지를 질적 평가한 선행연구 결과와 비교하였다. 연구결과 영향력 있는 논문을 판별하는 모델은 실제 수학교육 연구자들이 연구의 영향력을 평가하는 인식과 유사하게 설계된 것을 확인할 수 있었다.

지금까지 7개의 기계학습 분류모델 중 어떤 분류모델을 선택하면 수학교육 분야의 영향력 있는 논문을 판별할 수 있는지 확인하였다. 그 결과 랜덤포레스트 모델이 가장 우수한 성능을 보이고 있었으며, 수학교육 연구자의 질적 평가와 유사한 결과를 보이는 것을 확인할 수 있었다.

IV. 결론 및 제언

이 연구에서는 기계학습으로 수학교육 연구에 대한 메타분석을 실시하였다. 과학계량학의 관점에서 수학교육 논문의 영향력을 평가하는 모델을 설계하였다. 이 논문의 궁극적인 목적은 과학계량학의 관점에 따라 영향력 있는 수학교육 연구를 판별하는 것이었다. 기계학습으로 분류모델을 설계하는 선행연구를 분석하였다. 수학 학업성취도 예측모델을 설계한 안도연, 이광호(2022)의 연구는 설계모델을 랜덤포레스트로 확정하고, 다른 모델과 성능

비교 없이 랜덤포레스트 모델을 최적화하는 과정만 수행하였다. 마찬가지로 생명공학 분야의 영향력 있는 논문을 예측하는 기계학습 모델을 설계한 Weis, Jacobson(2021)의 연구도 문헌연구에 근거하여 설계모델을 랜덤포레스트로 미리 확정하였다. 하지만 기계학습 분류모델은 로지스틱 회귀, KNN, 나이브 베이즈, 의사결정트리, 랜덤포레스트, lightGBM, LinearSVC 등 여러 가지가 있었다. 수학교육 기초학력 미달 비율을 예측하는 모델을 설계한 이종현, 조규락(2021)의 연구에서는 여러 가지 기계학습 모델을 비교하여 가장 성능이 높은 모델을 채택하였다. 이에 본 연구에서도 7개의 기계학습 분류모델에 수학교육 논문의 메타정보를 입력하여 그 성능을 비교하였다. 랜덤포레스트 모델의 정확도는 0.980, 정밀도는 0.726, 재현율은 0.860으로 측정되었으며, 다른 모델과 비교하여 각 항목에서 가장 높은 성능을 보이거나, 두 번째로 높은 성능을 보였다. 일부 평가지표에서 가장 높은 점수를 획득한 다른 모델들은 다른 평가지표에서는 최하에 가까운 점수를 받았기에 모든 지표에서 균형 있게 높은 점수를 획득한 랜덤포레스트를 최종 모델로 채택하였다.

설계한 모델에 수집했던 24,575편의 논문의 메타정보를 입력하여 수학교육의 영향력 있는 논문을 판별하였다. 그 결과 24,575편의 논문 중 1,017편의 논문이 영향력 있는 논문일 것으로 예측되었다. 영향력 있는 논문이 게재된 학술지를 조사한 결과, 영향력 있는 논문들의 학술지는 JRME 25.66%, ESM 22.12%, ZDM 8.46%, JMTE 5.8%, JMB 5.51%, IJSME 5.51%, PRIMUS 4.92%, MTL 4.23%, SSM 4.03%, FLM 3.24% 순으로 나타났다. 이는 수학교육 학술지를 질적으로 평가한 Williams, Leatham(2017)의 연구결과와 대부분 일치하였다. Williams, Leatham(2017)의 연구에서 JRME과 ESM은 수학교육 연구자들에게 'very high'라는 평가를 받았으며 ZDM, JMB, FLM, MTL, JMTE는 수학교육 연구자들에게 'high'라는 평가를 받았다.

설계한 기계학습 모델은 Williams, Leatham(2017)의 연구에서 'very high'와 'high' 평가를 받은 두 그룹 학술지의 차이를 정량적인 수치로 제시할 수도 있는 장점이 있다. 영향력 있는 논문 중 'very high' 그룹의 학술지에 속하는 비율은 평균 23.9%, 'high' 그룹 학술지에 속하는 비율은 평균 5.48%이다. 설계한 기계학습 모델은 개별 논문의 영향력도 판별하지만, 두 그룹 학술지의 영향력의 차이도 판별할 수 있다.

한편 2017년부터 2021년까지 수학교육 학술지의 인용횟수를 조사한 Yig(2022)의 연구에서는 학술지별 피인용횟수는 ESM(1561), IJSME(947), JMTE(721), JRME(563) 순이었다. 피인용횟수로 연구의 영향력을 평가한다면, IJSME와 JMTE는 JRME보다 높은 평가를 받을 것이다. 하지만 IJSME는 과학교육 논문으로부터 받은 피인용횟수까지 포함되어 있다. 한편 수학교육 논문만을 다루고 있는 JMTE도 JRME보다 피인용횟수가 높다. 최근 교사교육에 관심이 증가하여 JMTE의 피인용횟수가 증가하였다고 분석할 수는 있지만, 수학교육 연구자들이 학술지를 평가하는 인식과 괴리가 있다. 이러한 괴리는 과학계량학의 관점에서 피인용횟수로 논문과 학술지의 영향력을 평가할 때 발생하는 문제의 한 사례가 될 것이다. 이 문제를 해결하기 위한 대안 중 하나로 본 연구에서 설계한 모델이 될 수 있음을 확인하였다. 기계학습 모델을 통해서 확인한 결과 영향력 있는 논문의 학술지 비율은 JRME 25.66%, JMTE 5.8%, IJSME 5.51%로, JRME가 JMTE 혹은 IJSME보다 학술지의 영향력에 있어서 약 4배 정도 높음을 확인할 수 있었다.

수학교육 연구의 영향력을 판별할 때 수학교육 네트워크를 형성하고 기계학습 모델을 활용하는 것이 수학교육 전문가 의견과 일치하는 경향을 보였다. 다시 말해 수학교육 전문가들이 직접 논문을 읽고 질적으로 평가한 것과 유사한 결과를 기계학습 모델을 설계하여 구현할 수 있다는 시사점을 얻을 수 있다. 수학교육 연구의 영향력을 질적으로 평가하기 위해서는 대규모의 연구진 인원과 분석에 많은 시간이 필요하다. 또한, 일부 연구자들은 질적 평가 방법에 대해서 의문을 제기(Andrade-Molina et al., 2020)하기도 한다. 전문가 집단의 의견을 기준으로 한 학술지의 평가는 전문가들 사이의 잠재적인 불일치 혹은 전문가들이 편집위원으로 활동하고 있거나 가장 자주 발행하는 학술지에 높은 점수를 부여할 위험(Haensly et al., 2008; Williams & Leatham, 2017)이 있기 때문이다. 기계학습을 활용한다면 이런 주관이 배제되며 데이터를 바탕으로 객관적으로 결과를 도출할 수 있다.

연구의 영향력을 질적으로 평가하기 위해서는 전문가들이 직접 논문을 읽고, 합의된 기준을 마련하며, 이를

근거로 평가하는 과정에서 많은 검토진과 시간이 필요했다. 본 연구의 기계학습 모델로 수학교육 연구의 영향력을 평가한다면 적은 인원과 짧은 시간 안에 전문가들의 질적 평가와 유사한 결과를 얻을 수 있다는 점에서 의의가 있다.

이 연구는 영향력 있는 수학교육 연구를 과학계량학의 관점으로 네트워크 중심성 지수 상위 5% 이내 논문으로 정의하고, 이를 판별하는 기계학습 모델을 설계하였다. 과학계량학의 관점에서 설계한 모델의 성능을 분석하고 타당성을 검증하였다. 또한, 기계학습 모델의 결정을 분석하고 그 타당성을 확인하였다. 이 연구에서 독립변인을 계산할 때, 인용에 가중치를 두지 않고 횟수로 카운트하였다. 하지만 논문에서 인용되는 참고문헌의 기여도는 다르기 때문에 모든 인용이 1로 취급되는 것은 수학교육 학문 네트워크의 현실을 정확하게 반영하였다고 보기에는 한계가 있다. 이에 따른 후속연구를 위한 제언은 다음과 같다. 인공지능의 자연어처리와 감성 분석을 활용하여 참고문헌을 긍정적으로 인용했으면 1, 부정적으로 인용했으면 -1 등과 같이 세분화하여 가중치를 부여한 이후 이 연구의 모델을 활용한다면 이 연구의 결과를 확장할 수 있을 것이다.

한편 이 연구에서 인공지능 모델에 입력한 독립변인의 개수는 140개이며 이를 연구자가 준비하기 위해서는 많은 시간이 필요하다. 이에 독립변인의 수를 축소할 수 있다면 더 효율적인 모델을 설계할 수 있을 것이다. 하지만 독립변인의 개수가 줄어들면 인공지능 모델의 정확도는 감소하기 때문에, 논문의 영향력을 판별하는 인공지능 모델에 독립변인의 영향력을 분석할 수 있는 인공지능 모델을 추가 설계한다면 이 연구의 결과를 확장할 수 있을 것이다.

참 고 문 헌

- 권선영 (2014). 구조적 학술용어사전 데이터베이스 구축에 있어서 용어의 의미관계 형성에 영향을 미치는 요인에 관한 연구. 한국문헌정보학회지, **48(2)**, 353-378.
- Kwon, S. (2014). A study on the factors influencing semantic relation in building a structured glossary. *Journal of the Korean Society for Library and Information Science*, **48(2)**, 353-378.
- 권오남 · 주미경 (2003). 대학 수학교육 연구의 동향과 과제. 수학교육, **42(2)**, 229-245.
- Kwon, O., & Ju, M. (2003). Research Trends and Tasks in Collegiate Mathematics Education. *The Mathematical Education*, **42(2)**, 229-245.
- 김관준 · 이재운 (2010). 학술지 영향력 측정을 위한 h-지수의 응용에 관한 연구. 정보관리학회지, **27(1)**, 269-287.
- Kim, P., & Lee, J. (2010). A study on journal impact measurement with hirsch-type indices. *Journal of the Korean Society for Information Management*, **27(1)**, 269-287.
- 문성구 · 김인재 (2018). 공동연구 네트워크에서 중심성과 연구성과 간의 관련성에 관한 연구. 소프트웨어 및 데이터 공학, **7(5)**, 169-176.
- Moon, S., & Kim, I. (2018). A study of the relationship between centrality and research performance in collaborative research network. *KIPS Transactions on Software and Data Engineering*, **7(5)**, 169-176.
- 박경미 (2003). 중등 수학교육 연구의 경향 분석 <수학교육> 과 <JRME> 에 수록된 논문 비교를 중심으로. 수학교육, **42(2)**, 219-228.
- Park, K. (2003). An Analysis of the Secondary Mathematics Education Research Trends in Korea by Comparing the Papers on and Those on. *The Mathematical Education*, **42(2)**, 219-228.
- 박선영 · 김원경 (2011). 국내외 수학교육 연구 동향 비교 분석. 수학교육, **50(3)**, 285-308.
- Park, S., & Kim, W. (2011). A comparative analysis on research trends of secondary mathematics education between

- Korea and overseas. *The Mathematical Education*, **50(3)**, 285 - 308.
- 방정숙 · 선우진 · 조선미 · 이유진 · 김은경 · 김윤영 · 박예진 · 김경훈 · 황지남 · 이하늬 (2019). 국내 수학교육 연구의 동향 분석-1963년부터 2019년까지 게재된 국내 수학교육 학술지 논문을 중심으로. *수학교육학 연구*, **29(4)**, 709-739.
- Pang, J., Jin, W., Cho, S., Lee, Y., Kim, E., Kim, Y., Park, Y., Kim, K., Hwang, J., & Lee, H.(2019). Domestic research trends of mathematics education: An analysis of journals published from 1963 to 2019. *Journal of Educational Research in Mathematics*, **29(4)**, 709-739.
- 안도연 · 이광호 (2022). 학업성취도 예측 요인 분석 및 인공지능 예측 모델 개발-블렌디드 수학 수업을 중심으로. *수학교육*, **61(2)**, 257-271.
- Ahn, D., & Lee, K. (2022). Analysis of achievement predictive factors and predictive AI model development - Focused on blended math classes. *The Mathematical Education*, **61(2)**, 257 - 271.
- 이재윤 (2008). 연구자의 투고 학술지 현황에 근거한 국내 학문분야 네트워크 분석. *정보관리학회지*, **25(4)**, 327-345.
- Lee, J. (2008). Analyzing the network of academic disciplines with journal contributions of Korean researchers. *Journal of the Korean Society for Information Management*, **25(4)**, 327 - 345.
- 이재윤 (2011). 국내 인용 데이터베이스에서 저널 페이지랭크 측정 방안. *한국비블리아학회지*, **22(4)**, 361-379.
- Lee, J. (2011). Journal pagerank calculation in the Korean science citation database. *Journal of the Korean BIBLIA Society for Library and Information Science*, **22(4)**, 361 - 379.
- 이재윤 (2014). 공동연구 네트워크 분석을 위한 중심성 지수에 대한 비교 연구. *정보관리학회지*, **31(3)**, 153-179.
- Lee, J. (2014). A comparative study on the centrality measures for analyzing research collaboration networks. *Journal of the Korean Society for Information Management*, **31(3)**, 153 - 179.
- 이정연 · 이재윤 · 강인수 · 신숙경 · 정한민 (2007). 확률적 온톨로지와 연구자 네트워크를 이용한 심사자 자동 추천에 관한 연구. *정보관리학회지*, **24(3)**, 43-65.
- Lee, J., Lee, J., Kang, I., Shin, S., & Jung, H. (2007). Automatic recommendation of panel pool using a probabilistic ontology and researcher networks. *Journal of the Korean Society for Information Management*, **24(3)**, 43 - 65.
- 이종현 · 조규락 (2021). 머신러닝을 활용한 중학교 수학 기초학력 미달 비율 예측모형 탐구. *교육공학연구*, **37(1)**, 95-129.
- Lee, J., & Cho, K. (2021). A study on the prediction model for the ratio of mathematics low-performing students in middle school using machine learning. *Journal of Educational Technology*, **37(1)**, 95-129.
- 조은성 · 송재도. (2013). 어떤 국내 논문이 더 많이 인용되는가. *마케팅연구*, **28(4)**, 33-56.
- Cho E., & Song J. (2013). 어떤 국내 논문이 더 많이 인용되는가?. *Korean Journal of Marketing*, **28(4)**, 33-56.
- Acuna, D. E., Allesina, S., & Kording, K. P. (2012). Predicting scientific success. *Nature*, **489(7415)**, 201-202.
- Aljuaid, H., Iftikhar, R., Ahmad, S., Asif, M., & Afzal, M. T. (2021). Important citation identification using sentiment analysis of in-text citations. *Telematics and Informatics*, **56**, 101492.
- Altuntas, S., Dereli, T., & Kusiak, A. (2015). Forecasting technology success based on patent data. *Technological Forecasting and Social Change*, **96**, 202-214.
- Andrade-Molina, M., Montecino, A., & Aguilar, M. S. (2020). Beyond quality metrics: defying journal rankings as the philosopher's stone of mathematics education research. *Educational Studies in Mathematics*, **103(3)**, 359-374.
- Barnes, J. A., & Harary, F. (1983). Graph theory in network analysis. *Social networks*, **5(2)**, 235-244.
- Bildosola, I., Río-Bélver, R. M., Garechana, G., & Cilleruelo, E. (2017). TeknoRoadmap, an approach for depicting emerging technologies. *Technological Forecasting and Social Change*, **117**, 25-37.

- Bollen, J., Rodriguez, M. A., & Van de Sompel, H. (2006). Journal status. *Scientometrics*, **69**(3), 669-687.
- Borrmann, L., Mutz, R., Neuhaus, C., & Daniel, H. D. (2008). Citation counts for research evaluation: standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in science and environmental politics*, **8**(1), 93-102.
- Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems*, **30**(1-7), 107-117.
- Chen, P., Xie, H., Maslov, S., & Redner, S. (2007). Finding scientific gems with Google's PageRank algorithm. *Journal of Informetrics*, **1**(1), 8-15.
- Ding, Y., & Cronin, B. (2011). Popular and/or prestigious? Measures of scholarly esteem. *Information Processing & Management*, **47**(1), 80-96.
- Dong, Y., Johnson, R. A., & Chawla, N. V. (2015, February). Will this paper increase your h-index? Scientific impact prediction. In *Proceedings of the eighth ACM international conference on web search and data mining* (pp. 149-158).
- English, L. D., & Kirshner, D. (Eds.). (2002). *Handbook of international research in mathematics education*. Lawrence Erlbaum.
- Forgasz, H. J. (2019). *Scholarly writing*. In *Compendium for Early Career Researchers in Mathematics Education*, 359-374. Springer.
- Fujigaki, Y. (1998). Filling the gap between discussions on science and scientists' everyday activities: Applying the autopoiesis system theory to scientific knowledge. *Social Science Information*, **37**(1), 5-22.
- Funk, R. J., & Owen-Smith, J. (2017). A dynamic network measure of technological change. *Management science*, **63**(3), 791-817.
- Garfield, E., & Sher, I. H. (1963). New factors in the evaluation of scientific literature through citation indexing. *American documentation*, **14**(3), 195-201.
- Garfield, E. (2009). From the science of science to Scientometrics visualizing the history of science with HistCite software. *Journal of Informetrics*, **3**(3), 173-179.
- Grouws, D. (Ed.). (2006). *Handbook of research on mathematics teaching and learning: (A project of the national council of teachers of mathematics)*. IAP.
- Grouws, D. A. (1988). *Perspectives on Research on Effective Mathematics Teaching. Volume 1*. National Council of Teachers of Mathematics, 1906 Association Drive, Reston, VA 22091.
- Grouws, D. A. (Ed.). (1992). *Handbook of research on mathematics teaching and learning*. MacMillan.
- Haensly, P. J., Hodges, P. E., & Davenport, S. A. (2008). Acceptance rates and journal quality: An analysis of economics and finance. *Journal of Business & Finance Librarianship*, **14**(1), 2 - 31.
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National academy of Sciences*, **102**(46), 16569-16572.
- Hood, W., & Wilson, C. (2001). The literature of bibliometrics, scientometrics, and informetrics. *Scientometrics*, **52**(2), 291-314.
- Inglis, M., & Foster, C. (2018). Five decades of mathematics education research. *Journal for Research in Mathematics Education*, **49**(4), 462-500.
- Johnson, W. R. (1987). Empowering practitioners: Holmes, Carnegie, and the lessons of history. *History of Education Quarterly*, **27**(2), 221-240.

- Lakatos, I. (1978). *The methodology of scientific research programmes: Philosophical papers Volume*. Cambridge, United Kingdom: Cambridge University Press. doi:10.1017/CBO9780511621123.
- Lee, H. J., & Yang, H. Y. (2015). 시간적 효과가 고려된 PageRank 함수를 이용한 핵심특허 탐색방법. In Proceedings of the Korea Technology Innovation Society Conference, 275-281.
- Lester, F. K. (2005). On the theoretical, conceptual, and philosophical foundations for research in mathematics education. *ZDM*, **37(6)**, 457 - 467. doi:10.1007/BF02655854
- Mariani, M. S., Medo, M., & Zhang, Y. C. (2016). Identification of milestone papers through time-balanced network centrality. *Journal of Informetrics*, **10(4)**, 1207-1223.
- Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica: Journal of the Econometric Society*, 867-887.
- Nalimov, V. V., & Mulchenko, Z. M. (1971). *Measurement of Science*. Study of the Development of Science as an Information Process.
- Neylon, C., & Wu, S. (2009). level metrics and the evolution of scientific impact. *PLoS biology*, **7(11)**, e1000242.
- Nivens, R. A., & Otten, S. (2017). Assessing journal quality in mathematics education. *Journal for Research in Mathematics Education*, **48(4)**, 348 - 368.
- Presmeg, N., & Kilpatrick, J. (2019). Pleasures, power, and pitfalls of writing up mathematics education research. *Compendium for early career researchers in mathematics education*, 347-358.
- Price, D. (1963). *Little science, big science... and beyond* (Vol. 480). New York: Columbia University Press.
- Price, D. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American society for Information Science*, **27(5)**, 292-306.
- Schoenfeld, A. H. (2007). Method. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 1, pp. 69 - 107). Information Age
- Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. *Bmj*, **314(7079)**, 497.
- Sierpinska, A., & Kilpatrick, J. (1998). *The ICMI Study Conference. In Mathematics Education as a Research Domain: A Search for Identity* (pp. 3-32). Springer.
- Walker, D., Xie, H., Yan, K. K., & Maslov, S. (2007). Ranking scientific publications using a model of network traffic. *Journal of Statistical Mechanics: Theory and Experiment*, **2007(06)**, P06010.
- Waltman, L., & Yan, E. (2014). *PageRank-related methods for analyzing citation networks*. In Measuring scholarly impact (pp. 83-100). Springer.
- Weihls, L., & Etzioni, O. (2017, June). Learning to predict citation-based impact measures. In *2017 ACM/IEEE joint conference on digital libraries (JCDL)* (pp. 1-10). IEEE.
- Weis, J. W., & Jacobson, J. M. (2021). Learning on knowledge graph dynamics provides an early warning of impactful research. *Nature Biotechnology*, **39(10)**, 1300-1307.
- William, D. (2003). *The impact of educational research on mathematics education*. In Second international handbook of mathematics education (pp. 471-490)
- Williams, S. R., & Leatham, K. R. (2017). Journal quality in mathematics education. *Journal for Research in Mathematics Education*, **48(4)**, 369 - 396.
- Wilson, C. (2001), Informetrics. In: M. E. WILLIAMS, (Ed.), *Annual Review of Information Science and Technology*, Vol.34, Medford, NJ: Information Today, Inc. for the American Society for Information

Science, 3 - 143.

Yig, K. G. (2022). Research trends in mathematics education: A quantitative content analysis of major journals 2017-2021. *Journal of Pedagogical Research*, **6(3)**, 137-153.

Zhou, Y., Li, Q., Yang, X., & Cheng, H. (2021). Predicting the popularity of scientific publications by an age-based diffusion model. *Journal of Informetrics*, **15(4)**, 101177.

Zhu, X., Turney, P., Lemire, D., & Vellino, A. (2015). Measuring academic influence: Not all citations are equal. *Journal of the Association for Information Science and Technology*, **66(2)**, 408-427.

Development of an impact Identification Program in Mathematical Education Research Using Machine Learning and Network

Oh, Se Jun

Ewha Womans University High School

E-mail : skyjune@sen.go.kr

Kwon, Oh Nam[†]

Seoul National University

E-mail: onkwon@snu.ac.kr

This study presents a machine learning program designed to identify impactful papers in the field of mathematics education. To achieve this objective, we examined the impact of papers from a scientific econometrics perspective, developed a mathematics education research network, and defined the impact of mathematics education research using PageRank, a network centrality index. We developed a machine learning model to determine the impact of mathematics education research and identified the journals with the highest percentage of impactful articles to be the Journal for Research in Mathematics Education (25.66%), Educational Studies in Mathematics (22.12%), Zentralblatt für Didaktik der Mathematik (8.46%), Journal of Mathematics Teacher Education (5.8%), and Journal of Mathematical Behaviour (5.51%). The results of the machine learning program were similar to the findings of previous studies that were read and evaluated qualitatively by experts in mathematics education. Significantly, the AI-assisted impact evaluation of mathematics education research, which typically requires significant human resources and time, was carried out efficiently in this study.

* 2000 Mathematics Subject Classification : 97B10

* Key words : mathematical education research, network, impact, machine learning

[†] corresponding author