

딥러닝 기반의 기술가치평가와 평가변수 추정

Deep Learning-based Technology Valuation and Variables Estimation

성태응, 김민승, 이찬호, 최지혜, 장용주, 이정희
연세대학교(미래)

Tae-Eung Sung(tesung@yonsei.ac.kr), Min-Seung Kim(kms903@yonsei.ac.kr),
Chan-Ho Lee(chpat2785@yonsei.ac.kr), Ji-Hye Choi(toddlf58@yonsei.ac.kr),
Yong-Ju Jang(lolhi@yonsei.ac.kr), Jeong-Hee Lee(smartbio@naver.com)

요약

국내 산업성장의 동력이 되는 기업의 기술 및 사업역량 확보를 위해, 2014년 이후 정부기관을 중심으로 『기술거래 시장 활성화』, 『기술금융 기반의 R&D 사업화지원』 등 다양한 형태로 사업화 성과창출을 위한 정책 프로그램이 수행되어 왔다. 현재까지 기술가치평가 이론과 평가변수에 관한 다양한 연구가 각계 전문가들에 의해 수행되고 기술사업화 현업에서 이용되어 왔으나, 평가현장에서는 평가목적 상의 기대보다 기술가치평가 금액이 적게 산출되는 등의 한계점이 존재하여, 그 실용성에 의문이 제기되어 왔다. 또한 미래에 대한 정성적 추정에 기반한 현행 기술가치평가 프로세스는 평가결과에 대한 객관성과 신뢰성 확보를 위해 데이터에 근거한 참조 인프라 구축이 필요한 시기에 이르렀다. 본 연구에서는 기관별로 구축된 평가인프라를 살펴보고, 최신의 인공지능 및 딥러닝 기술을 기술가치평가 시스템에 탑재하기 위해 핵심변수 기반의 가치금액 추정 시뮬레이션과 매출액 및 정성평가점수의 예측을 위한 적용 가능성을 살펴본다.

■ 중심어 : | 기술가치평가 | 딥러닝 | 인공지능 | 기술평가시스템 | 매출추정 | 정성평가지표 |

Abstract

For securing technology and business competences of companies that is the engine of domestic industrial growth, government-supported policy programs for the creation of commercialization results in various forms such as 『Technology Transaction Market Vitalization』 and 『Technology Finance-based R&D Commercialization Support』 have been carried out since 2014. So far, various studies on technology valuation theories and evaluation variables have been formalized by experts from various fields, and have been utilized in the field of technology commercialization. However, Their practicality has been questioned due to the existing constraint that valuation results are assessed lower than the expectation in the evaluation sector. Even considering that the evaluation results may differ depending on factors such as the corporate situation and investment environment, it is necessary to establish a reference infrastructure to secure the objectivity and reliability of the technology valuation results. In this study, we investigate the evaluation infrastructure built by each institution and examine whether the latest artificial neural networks and deep learning technologies are applicable for performing predictive simulation of technology values based on principal variables, and predicting sales estimates and qualitative evaluation scores in order to embed onto the technology valuation system.

■ keyword : | Technology Valuation | Deep Learning | Artificial Intelligence | Technology Assessment System | Sales Prediction | Qualitative Evaluation Indicators |

I. 서론

1. 국내 기술평가 현황

산업통상자원부, 과학기술정보통신부, 중소벤처기업부, 특허청, 기술보증기금과 같은 국내 정부부처 및 산하기관들은 기술사업화 자금지원이나 상장평가 등을 수행하기 위해 다양한 기술평가 프로그램을 운영 중에 있다.

한국산업기술진흥원(KIAT: Korea Institute for Advancement of Technology)가 공개한 『기술금융 지원사업 종합성과분석』결과에 의하면[1], 기술평가 결과는 특허청 및 금융위, 과학기술정보통신부, 한국거래소 등을 중심으로 사업화 및 투자 연계형 특허기술평가 지원사업, 기술신용대출사업, 상장특례평가사업 및 공공R&D 기술의 사업화 성과창출을 위한 기술가치평가 지원사업 등 다양한 유형으로 활용되고 있다. 정부의 기술 기반 금융활동 지원이 지속적으로 수행되고 있으나, 기술금융 산업 활성화의 핵심적 요인 중 하나인 신뢰성 있는 기술평가 도출을 위한 효과적인 인프라 구축 또는 객관성 검증을 위한 참조정보 체계가 현실적으로 구축되지 않은 실정이다.

현재 국내 평가전문기관이 수행하고 있는 기술평가 방안은 대부분 기술 자체가 보유한 내재적 역량을 중심으로 하며, 실제 사업화를 수행할 기업에 대한 정성적 평가가 주가 되고 있다. 한국발명진흥회(KIPA: Korea Invention Promotion Association)가 지정·시행하는 '발명의 평가기관' 제도는 투자, IP담보, 사업화 등의 목적성을 지니며, 평가대상 특허지표에 대한 상대적 우위성 분석, 특허 정량비교 분석, 사업타당성 및 가치평가 등을 수행한다. 동 사업은 2020년 기준 총 53억여 원의 예산을 투입하는 등 적극적인 기술금융지원을 수행 중에 있다.

2014년 이후 금융위원회가 운영하는 기술신용평가

시스템(Technology Credit Bureau, TCB)의 경우, 기술보증기금과 4대 신용평가기관(NICE평가정보, 한국기업데이터, 이크레더블, 서울신용정보)를 중심으로 연간 150~200여 건씩 평가가 수행되고 있다. 기존 TCB 체제는 재무구조에 기반한 신용등급평가를 주로 수행해 왔으나, 재무적 역량이 다소 미흡한 중소·벤처기업의 특성을 고려하여 기업이 보유한 기술력에 대한 평가를 추가 적용함으로써 종합적으로 기술신용등급을 최종 산정하는 평가방식으로 보완 적용되고 있는 추세이다[2].

2. 기술가치평가 관련 연구개발 현황

과학기술정보통신부에서 수행 중에 있는 기술가치평가사업은 대학과 출연연이 보유한 특허 등 공공 R&D 성과에 대해 정량적 가치평가를 적용함으로써, 기술이전 및 거래, 보증 등에 대한 근거자료로 활용되는 등 기술금융시장 전반에 대한 활성화에 기여하고 있다.

[표 1]에 의하면, 2016년부터 2020년 상반기까지 총 3만 9천여 건(누적)의 기술평가가 수행되었는데, 2017년 1만 4천여 건을 기점으로 지속적인 감소 추이를 보이고 있음을 확인하였다. 평가 유형별로는 2018년까지 기술력등급평가의 비중이 절대적으로 높았으나, 2019년에 들어서 기술력평가보다 기술가치평가의 건수가 더 증가하였는데, 일반적으로 정량적인 금액 기반 산출이 가능한 기술가치 평가건수의 증가는 향후 기술이전·거래 및 보증 등 정량적 금액에 기반한 기술금융 활동의 확산가능성과 이를 활용한 사업화 성과창출이 기대되는 대목이다.

그러나 실무에서의 기술가치평가에 대한 수요가 증대됨과 동시에 연구개발서비스업에 종사하는 전문평가 인력의 수가 늘어나고 있는 추세에서, 기술 및 사업 주체의 재무역량에 대한 정성적 평가를 위주로 수행되는 현행 기술가치평가 체계에서의 객관성 및 신뢰성 확보

표 1. 2016~2020 기술평가사업 지원 현황

	2016	2017	2018	2019	2020(상반기)	합계
기술가치평가	636	1,673	1,696	2,172	1,428	201
기술력등급평가	12,694	12,196	5,851	904	201	31,846
R&D 경제성평가	7	113	47	41	-	208
합계	13,337	13,982	7,594	3,117	1,629	39,659

는 여전히 미흡한 실정이다.

현재 실무에 적용되고 있는 기술가치평가 프로세스는 대부분 전문가의 정성적 평가에 의존하여 주관적인 관점이 개입되는 가치금액 산출이 주를 이루고 있어, 평가근거에 대한 신뢰성에 다소 제약이 존재한다. 또한 연구개발서비스 분야에 종사하는 기술이전 사업화 전담조직(TLO)이나 비전문가가 기술가치평가를 수행할 경우, 타 평가기관의 가치평가 결과가 외부에 공표되지 않으므로 기술평가 실무가이드[3]와 같은 매뉴얼 지침서나 STAR-Value 시스템[4]와 같은 웹 기반 지능형 평가시스템을 제외하고는 평가에 참조 가능한 정보가 다소 제한적인 실정이다.

기술가치평가의 활용성과 보급을 확대시키기 위해서는 평가결과를 이용하는 수요자(벤처투자자, 여신기관, 기술이전 매매 의향자, Licensee 등)에게 결과의 고신뢰성을 확보해야 한다. 이러한 기술가치평가 기법의 검증은 위해서는 과거 실제 평가된 기술에 대해 현 시점까지 발생한 사업가치를 추적조사한 후, 과거 기술가치평가 결과와 비교·분석하는 추적조사를 적용할 수 있으며 이를 통해 예측에 기반한 기술가치평가 결과와 향후 창출 가능한 사업가치 기반 평가액의 신뢰성을 고도화할 수 있을 것으로 예상된다. 그러나 상기의 방법론은 기술(특허)라는 무형자산이 사업에 대한 경제적 기여분을 고려하여 미래 시점에도 명확히 산출할 것이라는 긍정적 전망이 전제된다는 점과 각 평가건에 대한 추적조사 자체가 내포하고 있는 비효율성 등으로 인해 실제적인 검증 적용에 한계가 있을 것으로 예상된다.

최신 기술로 부각된 인공지능 및 빅데이터 기술이 아직 기술가치평가에 실질적으로 적용되어 운영되는 사례는 기술보증기금(KIBO)의 KPAS 시스템 외 전무한 실정이다. 기술가치평가 산출로직 상의 기술의 경제적 수명 및 현금흐름, 할인율, 기술기여도 등 다양한 평가요인(evaluation factors)이 존재함을 고려하였을 때, 평가대상기술 및 사업주체에 대한 데이터를 확보 가능할 경우, 인공지능 기반 모델링을 적용하여 평가요인별 추정값을 도출해 내는 것이 가능할 것으로 여겨진다. 특히 인공지능 기반의 기술가치평가 방식이 적용될 경우, 재무데이터 또는 평가사례정보 등의 증거(evidence) 기반 가치평가결과를 반영하여 신속하게

기술가치평가 범위(또는 값)를 추론할 수 있을 것이며, 사업화를 목적으로 하는 다양한 비전문가 집단(TLO담당, 평가기관 내 기술평가담당, 민간 사업화지원 컨설턴트 등)에게도 손쉽게 접근가능한 편리한 분석 툴(tool)이 될 것으로 기대된다. 동시에 기술가치 평가결과를 축적하여 DB화함으로써, 평가변수의 요인분석을 수행하거나 기술이전·현물출자 등 용도별 평가활용사례 DB 구축을 통해 필요시 참조하도록 할 수 있다. 이에 부가하여, 기술가치 평가체계에 대한 표준화 및 유사기술에 대한 신뢰성 있는 가치평가 도출범위를 제시하며 과거 평가사례결과의 DB화 및 핵심 변수별 참조 메타 DB를 통해 평가신뢰성에 대한 근거정보를 제공할 수 있을 것으로 기대된다.

따라서 본 연구에서는 현재 실무에 적용되고 있는 기술가치평가 방법론에서의 주요 평가요인별 적용 가능한 딥러닝 기법의 탐색을 통해 기술가치평가에서의 인공지능 및 빅데이터에 대한 적용 방향성을 제시하는 것에 주안점을 두었다.

3. 인공지능과 기술가치평가

글로벌 시장분석 조사기관인 Gartner의 데이비드 설리(David Cearley)는 인공지능(Artificial Intelligence, AI)이 메타버스, 증강현실/가상현실(AR/VR), 사물인터넷(IoT), 양자컴퓨팅, 디지털 트윈 등과 함께 이용되어, 궁극적으로 어플리케이션의 기능적 측면과 비기능적 측면을 모두 자동화하는 매우 고도화된 인공지능 주도 개발 환경이 비전문가들도 코딩 없이 AI 프레임워크를 활용한 어플리케이션을 만들 수 있게 해주는 혁신 수준의 AI 유연성을 제공할 것이라고 언급하였다[5].

인공지능은 일반적으로 컴퓨터공학 분야에서 분류 및 예측 등의 문제 해결을 위해 시작되었고 인간을 대신하는 컴퓨터의 자동화 알고리즘에 대해 다양한 이론이 파생되었으며, 최근에는 CNN(Convolutional Neural Networks) 및 RNN(Recurrent Neural Networks), LSTM(Long Short-Term Memory) 및 GAN(Generative Adversarial Networks) 등의 코어 소스코드가 오픈소스로 공개됨에 따라 비전문가들도 해당 소스코드를 분석 대상에 접목하고 최적화하여 인공지능의 적용 분야를 광범위하게 확장시켜 왔다.

예를 들어, 국방 분야에서의 인공지능 적용은 지능형 실시간 M&S 및 위게임체계, AI 기술을 적용한 지능형 표적탐지레이더, AI 사이버 방호시스템 등이 존재하며, 의료분야에서는 정밀의료 솔루션인 Dr. Answer 및 영상의학용 초음파 진단기기인 S-Detect, 뇌 영역 의료 영상분석을 수행하는 뷰노메드 딥브레인이 산업에 적용 중에 있다. 실생활에서 일반적으로 접할 수 있는 인공지능의 예로는 빅스비 및 기가지니, 챗봇 등이 존재하는 등 소 산업 영역에서 인공지능(기계학습, 딥러닝)은 다양한 목적성을 가지고 적용되고 있는 것으로 분석되었다.

인공지능의 개념은 1956년 존 매카시 교수가 처음 사용하는데 유래하며, 최종적으로 인간의 지능을 기계(컴퓨터)로 구현, 제작하는 것이 목표이다. 우리가 기대하는 수준의 인공지능은 마치 인간과 동일하게 스스로 이해하고 판단하는 ‘강 인공지능(General AI)’의 형태이지만, 현재의 기술 수준에서 구현 가능한 인공지능은 특정 작업에 대한 수행을 목표로 학습에 기반한 ‘약 인공지능(Narrow AI)’의 형태를 의미한다[6].

Mr Collins Belly	Attention Funds Owner, - Attention Funds Own
Anti-Fraud Unit ICPC	PAYMENT NOTIFICATION - ICPC NIGERIA (An
Zipeem	Per raggiungere i tuoi progetti, scopri le rego
Anti-Fraud Unit ICPC	PAYMENT NOTIFICATION - ICPC NIGERIA (An
UsTrendy indie Fashion	New Arrivals Just Added! - This message was
YOU'RE-APPROVED	YOUR Roof is Covered FREE!!
WORK @ HOME	(PLEASE REPLY) Woburn FULL TIME - 2 Rem
Office	Re:Confirm deposit - DEPOSITED INHERITAN
Mr.Sambo Ngene	Message From Mr.Sambo Ngene, Director, Forei
Congratulations!	Here is your chance to win a brand new Ford!
Jocelyn Weir	Do not regret skipping these Jobs - Jul 27, 20
Registered_Offender_List.	SEX-Offender living near-YOU!! (Find-Out)

그림 1. 머신러닝의 메일함 스팸 필터링 사례

※ 출처 : <https://blogs.nvidia.co.kr/>

머신러닝(기계학습, Machine Learning)은 인공지능의 하위 개념으로서, 의사결정 기준에 대한 구체적인 지침을 소프트웨어에 직접 코딩하는 것이 아닌, 대량의 데이터와 알고리즘을 통해 컴퓨터 그 자체를 학습시켜 작업수행 방법을 익히는 개념이며, 쉬운 예로 [그림 1]과 같이 메일함의 스팸(spam)을 자동으로 걸러내는 작업을 연상할 수 있다.

머신러닝의 하위개념으로서, 딥러닝(Deep Learning)은 완전한 머신러닝을 구현하는 자동화 알고리즘으로 간주할 수 있다. 딥러닝은 인공신경망으로부터 발전한

형태의 인공지능 알고리즘으로, 뇌의 뉴런과 유사한 정보 전달체계를 통해 데이터의 비선형적 패턴을 학습하며, 데이터를 인식하는 순간부터 예측이나 분류하는 순간까지 전 주기적으로 자동화된 판단을 수행하는 개념이다.

정지 표지판 사례의 경우, 색상 및 표시문자, 팔각형 모양 등과 같은 이미지의 특성이 추출되고 가중치가 적용되어 학습되며, 10,000개의 정지 표지판과 같은 이미지를 정확하게 인식하는 횟수가 9,200번일 경우, 모형의 정확도가 92%라고 설명한다. 신경망 모형이 최적화된다는 것은 수 만~수백 만 이미지를 학습함으로써 각 노드 간 가중치가 결합되는 과정을 의미한다. Tensorflow라는 고속 프레임워크를 개발한 구글과 앤드류 응(Andrew Ng) 교수는 1만 6천여 개의 컴퓨터로 약 10억 개 이상의 신경망으로 이뤄진 ‘심층신경망(Deep Neural Networks, DNN)’을 통해 컴퓨터가 고양이의 형태를 인식하고 스스로 판단하는 과정을 학습, 인간과 고양이의 사진을 자동으로 분류하는 모형을 구축하는 것에 성공한 바 있다.

그렇다면, 기존에 평가 대상 기술이 창출하게 될 시장과 사업주체의 예상점유율을 기반으로 재무적 가치를 산출하는데 기반하는 기술가치영역에서 인공지능(또는 딥러닝) 개념을 적용할 수 있는가? 즉, 기술가치 평가의 수익접근법 또는 로열티공제법 모형에도 심층신경망(DNN) 적용이 가능한 지에 대한 연구적 의문을 가져야할 시점이다. 평가사례가 수천~수만 건이 누적되고 데이터 완결성이 보장되는 일정 크기의 입출력 필드가 DB화되면, 다수의 은닉층으로 구성된 DNN을 최적화하여 블랙박스를 구성하여 학습을 통한 가중치 행렬(Weighted Matrix)가 형성되고, 신규 평가사례 건에 대한 변수정보들이 알려져 있고 그 값들이 입력될 경우 출력단에서의 가치평가액 결과예측에 활용할 수 있다.

본 연구에서는 최근까지 이용되어 온 특허 및 재무데이터 기반의 가치금액 산출방식으로부터 한층 발전시켜, 과거 평가사례와 유사한 사례가 입력될 경우 인공지능 기반 가장 그럴 듯한(highly probable and most feasible) 주요 평가요인 결과값의 추정 방안을 모색하고자 하였으며, 확보 가능한 평가사례를 기반으로

로 주요 평가요인에의 인공지능 추정을 적용한 결과 및 방향성을 제시하고자 하였다.

II. 연구내용

1. 인공지능경망과 기술가치평가

[표 1]에서 살펴본 바와 같이, 국내 공식적으로 집계되는 기술가치평가 건수는 2019년 기준 약 2,170여 건이며, 이러한 평가사례가 지속적으로 누적될 경우, 다양한 산업에서의 다각화된 기술평가 적용 사례가 확보될 것이다. 이는 즉 가치평가 변수 및 관련 지표, 가치평가액을 DB화하여 누적할 경우, 심층신경망(Deep Neural Networks, DNN) 등 인공지능 기반 모델링하기에 충분한 학습데이터 확보가 가능함을 의미한다.

특히 기술가치평가 시, 기술산업의 트렌드 및 기업의 재무상태가 지속적으로 변화함에 따라 평가결과에서의 세부항목별 중요도가 달라질 수 있음을 고려할 필요가 있다. 이는 즉 기술가치평가 프로세스에 대한 자동화 구축 시, 평가항목별 복잡한 비선형적 패턴에 대한 반영이 가능하여야 하며 데이터 누적에 따른 지속적인 업데이트가 수행될 필요가 있음을 의미한다. 인공지능경망의 경우, 초기 학습데이터에 대해 인공지능경망의 층별 노드간 가중치를 최적화하고 하이퍼파라미터를 구축하면, 신규 평가대상 변수들을 설명변수에 대입하고 블랙박스 가중치를 통과하여 출력값(가치평가액 예측값)을 도출할 수 있다. 다음 년도에 추가적으로 확보된 평가 사례를 포함하여 학습을 수행할 경우, 인공지능경망의 층별 노드간 관계가 변화하고 입출력 값들의 비선형적 패턴을 가장 잘 해석하는 가중치로 모형의 업데이트가 수행되어, 상기의 과정을 잘 설명하는 복수의 심층 신경망(DNN)은 이전에 비해 신규 누적된 전체 4,200여 건의 평가사례 입출력 관계를 가장 잘 나타내도록 최적화되며, 이는 변수간 특성(feature)을 기반으로 최적화된다.

상기와 같은 프로세스를 5~10여년 간에 걸쳐 반복 적용하고 대략 1만~2만여 건의 평가사례에 대한 학습이 수행될 경우, 최적화된 신경망 모형은 향후 적용될 신규 평가건에 대한 입력값들과 유사한 사례를 탐지해

내게 된다. 이후 사전에 다수 학습되었을 가능성이 높고 이로부터 기본 가치평가사례들 중 가장 유사한 형태의 가치평가 결과들의 가중 결합식(Weighted Aggregate Sum)으로 최종 평가결과를 도출할 수 있다.

다만, 도출된 신경망의 신뢰성 제고를 위해 학습 시 입력변수들 중 일부 데이터 필드가 비워진 경우, 보정 또는 제외의 방식으로 전처리 작업을 수행해야 하며, 학습에 활용되는 훈련 및 테스트 데이터의 완결성을 고려할 필요가 있다. 이러한 데이터 분석은 기존의 컴퓨터공학-산업공학 부문에서 관심을 두는 Data Scientist의 영역이지만, 기술가치평가의 전체 프로세스 및 정량·정성적 지표 활용부문을 잘 이해하여 Domain Knowledge를 충분히 숙지하고 있어야 가능한 부분이라 할 수 있다.

국내에서는 2014년경에 들어서서 한국과학기술정보연구원(Korea Institute of Science and Technology Information, KISTI), 기술보증기금 등의 기관을 중심으로 데이터 기반 지능형 가치평가시스템이 연구·개발되어 왔으며, 이미 많은 기술평가 전문기관·기업에서 사업화 지원을 위한 가치평가 예측에 해당 시스템을 활용하고 있다[7]. 이에 부가하여, 상기와 같은 지능형 가치평가시스템의 효율적 평가 및 고도화를 위한 효과적 데이터 셋의 구성과 딥러닝(Deep Learning) 기반의 사례학습 모형을 추가 적용하는 방향에 대한 연구는 거의 없는 실정이다.

이에 본 연구에서는 자체 확보한 기술가치평가 사례(162건)를 적용한 회귀분석 모형과 딥러닝 알고리즘의 예측 결과를 비교 분석하였다. 해당 실험은 기술가치평가 핵심지표인 할인율, 기술기여도, 매출추정액을 독립변수로, 기술가치 평가액을 종속변수로 적용하였으며 이에 기반한 다중회귀분석 수행 결과, [표 2]에서와 같이 일반적으로 사회과학 현상분석에서의 유의성 기준 대비 약 0.6로 높은 Adjusted-R2를 나타내고 있음을 확인하였다[8].

표 2. 기술가치평가 사례 162건에 대한 회귀분석 결과

Variable	Estimate	Adj. R-Sq
Intercept	-3.8805	0.5991
Sales	1.5227	
Discount rate	-1.3333	
Technology factor	0.7403	

이는 독립변수로 적용된 핵심지표들이 수식 (1)과 같은 현금흐름할인법(DCF) 기반 기술가치평가 결과에 주요한 영향을 미침을 감안하여, 높은 설명력을 나타내는 것으로 해석 가능하다.

$$Tech\ Value = \sum_{t=1}^n \frac{FCF_t}{(1+r)^t} \times Tech\ Factor \quad (1)$$

동 데이터를 활용하여 비선형적 패턴학습 기반 딥러닝 모델을 구축한 결과는 다음과 같다.

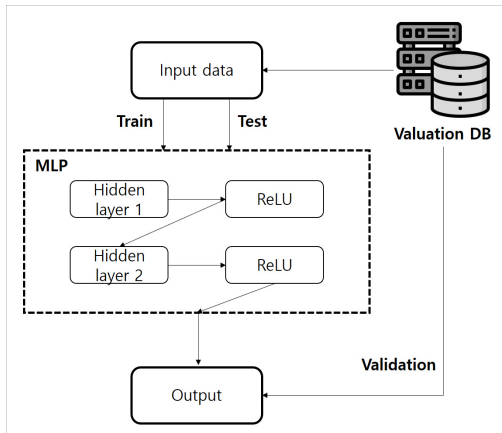


그림 2. 기술가치평가 사례 학습에 대한 MLP 모형의 흐름도

성능을 평가하는 주요 기준 중 하나인 MSE(Mean Squared Error)를 최소화시킴으로써, 최적화된 딥러닝 모형은 학습 데이터의 수가 상대적으로 적음을 감안하여, 2개의 hidden layer를 가지는 MLP (Multi-Layer Perceptron)의 형태로 적용되었으며 총 132개의 train data, 30개의 test data로 분할되어 200 Epochs, 0.01 learning rate를 기준으로 학습이 수행되었다. 최종적으로 도출된 딥러닝 기반 기술가치평가 모형의 평균예측력은 설정된 오차범위 threshold 기준 약 84%로 우수한 성능으로 판단되었으며, 향후 평가사례가 증대되고 모형의 하이퍼파라미터를 고도화시킴으로써 성능은 90%대로 높아질 수 있음이 확인되었다.

표 3. 적용 MLP 모형의 hyper-parameter 구성

구분	구성
Train : Test	132 : 30
Cost Function	ReLU
Framework / Model	Pytorch / MLP
Hidden layer /nodes	2 / 64, 128
Learning rate	0.01
Epochs	200

이러한 인공지능경망 기반의 가치평가 예측은 회귀분석결과와의 비교를 통한 변수별 유의성 분석, 연도별 추이에 따른 매출발생 기반 사업가치 추적조사 등의 검증들을 통해, 기술가치 평가결과와 객관성 및 신뢰성을 제고할 수 있을 것으로 기대된다.

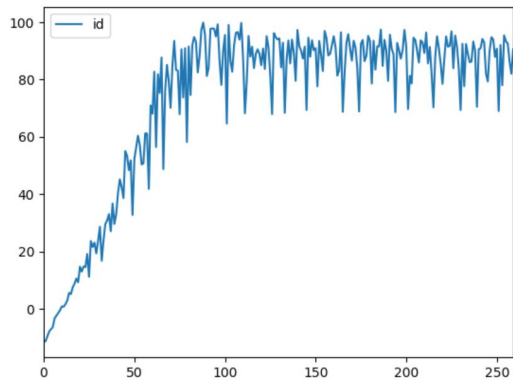


그림 3. 기술가치금액 예측모형에 대한 MLP 학습성능

2. 생성적 적대신경망(GAN)과 데이터 증강

앞선 딥러닝 기반 기술가치평가액 예측 모형은 부족한 학습데이터로 인해 보다 다양한 평가 패턴에 대한 학습에 한계가 존재한다. 적은 양의 학습데이터는 모형 구조의 단순화 및 학습에서의 과대·과소적합을 쉽게 유발할 가능성이 높아 딥러닝에서의 활용 근거를 다소 약화시키는 원인 중 하나이다[9][10]. 이러한 데이터 불균형(Data imbalance)의 문제는 SMOTE(Synthetic Minority Over-sampling TEchnique)와 같은 over-sampling기법의 적용을 통한 데이터의 증강 [11] 혹은 부스팅 등의 알고리즘을 적용한 학습데이터와 검증데이터 간 under-sampling을 통한 데이터의 불균형 해소[12]를 적용하는 등, 데이터 불균형을 극복

하고 심층신경망의 학습효율을 확보하는 방안이 다수 고안되고 있는 추세이다.

특히 2014년 인공지능 Top conference 중 하나인 NIPS(Neural Information Processing System)에서 알려진 생성적 적대 신경망(Generative Adversarial Networks)은 학습데이터를 기반으로 알고리즘으로부터 원 데이터와 생성데이터에 대한 구분을 기준으로 원 데이터와 상당히 유사한 패턴의 데이터를 신규 생성한다는 점에서 최근의 데이터 증강(Data Augmentation) 관련 연구에 활발히 적용되고 있다[13][14].

본 연구에서는 이러한 GAN 알고리즘 기반의 기술가치평가 데이터 증강이 현재 기술가치평가 사례데이터의 과부족을 일정 부분 완화해줄 수 있을 것이라 판단하여, 앞선 실험에 既 적용된 162건의 수집데이터에 대한 GAN 기반 데이터 증강을 적용하여 [그림 5]와 같은 생성사례 데이터의 기술가치 분포가 원 데이터와 유사한 분포를 띠을 확인하였다. 이는 확보된 최소한의 기술가치평가 사례를 기반으로 유사한 패턴의 가치평가 사례를 임의로 생성하여, 추후 심층신경망에의 학습 적용 시 데이터의 분포가 평가사례 모수의 분포와 유사한 형태를 나타내도록 하며 데이터 과부족의 보완적 요인으로 적용 가능함에 대한 분포 기반 적용 가능성 검증은 보여준다고 할 수 있다.

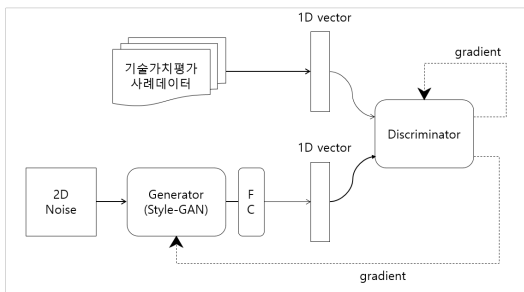


그림 4. 적용 GAN 알고리즘의 학습 흐름도

162건의 평가사례 데이터에 GAN 알고리즘을 적용한 결과, 생성된 가상 사례데이터는 실제 사례데이터가 내포하고 있는 가치금액 분포의 규모에 대한 특성은 유사하게 유지하되, 사례 부족으로 인해 통계적 패턴 확보에의 한계를 나타내었던 형태 대비 가상 데이터가 일

종의 Weibull distribution의 형태로 보완 적용되게 하였다. 이를 통해, 데이터의 사례가 실제 가치평가 시 기대되는 평가결과의 패턴과 유사하게 도출되는 것을 확인하였다[15]¹. 최근 2017~2019년 수행된 2천여 건의 기술가치평가 결과를 검토 분석한 결과, 실제 중소기업이 요청하는 기술가치평가 결과의 약 70% 이상이 3억 원에 가깝게 분포해 있다. 상기의 인공지능 기반 데이터 증강 기법은 기술가치평가와 같이 학습데이터의 과부족이 일반적으로 발생하는 학습사례에서 매우 효과적으로 적용 가능할 것으로 기대된다.

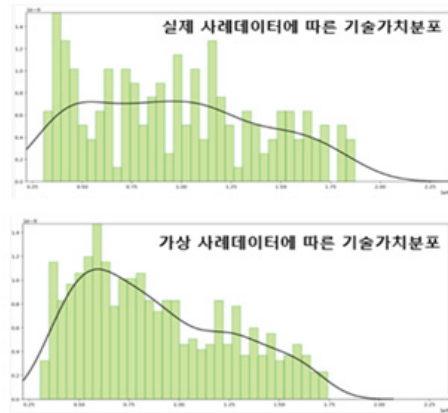


그림 5. GAN 알고리즘을 적용한 실제 사례데이터와 가상 생성사례 데이터의 기술가치 분포 비교

III. 연구결과

1. 딥러닝을 적용한 매출추정

기술사업화 추진 시 평가대상기술을 통해 발생 가능한 매출의 추정은 기술가치평가의 결과에 가장 중요한 영향을 미치는 요인 중 하나이며, 이러한 매출 추정은 사업 주체의 역량과 비즈니스 모델에 따라 달라지고, 최근까지도 주로 전문가가 직접 평가 및 추정하는 방식으로 수행되고 있다.

기업이나 업종별 재무정보를 제공하는 한국기업데이터나 나이스디앤비로부터 확보한 최근 10년간 업종별

1 가치평가 실무에서는 가치금액이 3억 부근에 집중적으로 분포하는 것이 일반적임.

매출액 정보 및 감쇄 패턴을 인공지능 기반으로 사전 학습시켰다. [그림 6]에서와 같이 과거 매출액 성장률을 연평균복합성장률(CAGR: Compound Annual Growth Rate)로 가정하고 동 성장률의 성장패턴을 낙관적, 중립적, 보수적으로 반영하여 평가대상 사업주체의 초기 3개년도 매출액이 주어질 경우, 이후 매출액 추정이 가능하다. 여기서, CAGR은 해당 업종의 기업유형(대/중/소/창업) 성장률에 따라 학습시킨 후 메타 DB를 사전 구축하고, 매출 추정 시 낙관적, 중립적, 보수적에 각각 매핑하여 [그림 6]와 같은 시나리오가 도출되도록 자동 적용 가능할 것으로 분석하였다[16].

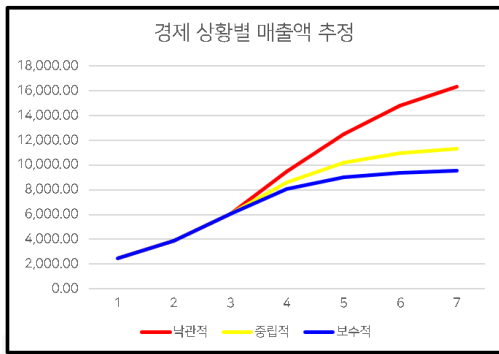


그림 6. 경제 상황별 시나리오에 근거한 매출액 추정 예

실제 기술가치평가 전문가들은 매출 추정 시 시장보고서를 기반으로 타겟시장의 형태 및 시장참여자의 구성, 시장의 확장가능성 등을 분석하여 사업주체의 예상 점유율을 추정하는 Top-down 방식을 적용하는 것이 일반적이다. 그러나 시장환경(경제상황), 투자계획 등에 따라 매출 추정 결과는 달라지므로, 전문가가 시간을 투입하고 정보를 분석하여 도출하게 되기까지는, 인공지능 기반의 매출 추정방식이 신속한 참조정보를 제공하는 것으로서 유용하다고 하겠다[17].

2. AI를 적용한 기술가치평가 정성지표 추정

기술가치평가에서의 AI(또는 딥러닝) 적용은 기술가치금액 외 다른 타겟변수(예. 정성평가 체크리스트 종합점수: 수명영향요인, 사업화위험프리미엄, 개별기술강도)에 대해 학습 결과를 반영함으로써 효율적 활용이

가능하다. 기술평가 실무가이드[2]에 의하면, 수익접근법 기반의 가치 산정 시, 기술수명 영향요인, 사업화 위험 프리미엄, 개별기술강도의 각 세부 항목별 평가는 여전히 전문가의 정성적 평가를 기반으로 수행되고 있으며, 이는 평가 결과에 대한 주관적 편향(bias) 및 신뢰성 저하를 야기하고 있는 대목 중 하나이다.

수익접근법과 로열티공제법의 경우, 매출액 추정 이후의 기술가치 산정 프로세스는 산식화 및 표준화를 통해 STAR-Value 등 지능형 가치평가시스템에 자동형으로 적용된 바 있다[2][7]. 그러나 동 시스템은 특허 및 재무정보 기반으로 현금흐름 산출을 위한 재무산식 기반으로 도출되고 있으며, 기술수명 영향요인 등 정성평가 지표(3개)는 여전히 전문가에 의해 평가되는 한계점이 존재한다.

표 4. 기술의 경제적 수명 정성평가 지표요인[3]

구분	세부요인	평 점				
		A	B	C	D	E
기술/권리요인	우월성					
	기술경쟁강도					
	대체가능성					
	모방난이도					
	권리보호강도					
시장/사업요인	시장진입 가능성					
	시장경쟁강도					
	시장경쟁의 변화					
	신제품 출현가능성					
	예상 시장점유율					
영향요인 평점 합계						

[표 4]와 같이 평가대상기술은 3개 정성평가 지표점수(기술의 경제적 수명 및 사업화 위험 프리미엄, 개별 기술강도)에 대해 전문가 평점이 주어지고[3], 이것이 수익예상기간, 할인율, 기술기여도에 각각 반영되어 최종 기술가치평가액을 산출하게 된다. 실제 평가사례가 수천 여 건 쌓여서 DB화 되었다고 할 경우, 해당 평가 사례가 내포하고 있는 특허, 기업개요정보, 사업주체의 비즈니스모델 업종정보 등을 설명변수로 하고 3개 지표점수를 타겟변수로 하는 딥러닝 모형을 구축한다면, 이에 대한 자동화된 추정이 가능할 것이다. 다만, 전술한 바와 같이 구축된 학습모형의 최적화를 통해, 훈련

데이터에 대한 오버피팅 가능성을 줄이고 테스트 데이터의 예측 성능을 극대화 시키는 것이 필요하다.

이 때, 딥러닝 기반 학습 적용을 위해 학습변수에 대한 구성은 기술평가 실무가이드에서 제공하는 정성평가 체크리스트를 특허데이터 및 사업주체의 재무데이터 항목과 매핑하고, 각 정성평가지표에 대한 평가특성을 원활히 반영할 수 있는 학습변수 pool을 각기 메타DB화하여 사전 구축해 둘 필요가 있다. 이와 같이 구축된 학습변수 메타DB는 평가자들이 오프라인 평가시 필요에 의해 부분적으로 정보 참조 가능하기도 하며, 누적된 평가사례데이터를 기반으로 업종별 평가건수 및 가치액 정보에 대한 통계 기능을 시스템화 할 수도 있다. 궁극적으로는 STAR-Value에서 제공하는 전문가의 심층 의견 입력 없이 현금흐름 산식으로만 계산 가능한 기본 기술가치결과를 토대로 하고, 3개 정성평가 점수가 딥러닝 모형으로부터 제시되어 반영될 경우 인공경망 기반의 기술가치 평가시스템이 평가 전문가 이외에 비전문가에게도 R&D 및 사업화 성과평가와 지원에 유용하게 활용될 것으로 기대된다.

IV. 결론 및 시사점

본 연구에서는 국내 기술가치평가 지원사업 현황과 인공지능(딥러닝)을 활용한 기술가치평가 매출추정 및 정성평가점수 예측에 대한 적용 가능성을 제시하였다. 동 연구는 기술가치평가의 핵심 변수요인으로부터 입력변수와 출력변수를 설정하고 다층의 딥뉴럴네트워크(DNN)를 구성하였으며, 既 확보된 기술가치 평가사례 데이터를 GAN 알고리즘 기반으로 충분히 증대시켰다. 또한, 추후 기술평가 누적사례가 증가할 경우 다양한 사례 기반의 고도화 된 학습을 통해 기술가치 범위(값) 추정 등에 확장 적용 가능할 것으로 분석하였다.

이외에도 기존 기술가치 평가사례로부터 기술수명 영향요인, 사업화 위험 프리미엄, 개별기술강도 점수가 라벨링되어 학습모형에 활용되고, 신규 평가대상 기술의 입력변수(대상 기술의 특허정보 및 기업정보, 기업 유형 등)에 대해 타겟변수인 정성평가지표를 예상하여, 모형의 성능을 높이는 최적화과정도 수행 가능한 것

로 판단되었다.

이상에서 살펴본 바와 같이, 전문가 평가 방식이던 기술평가체계를 특허 및 재무DB 기반으로 자동 산출하는 STAR-Value 시스템이 최근까지 자주 활용되고 있다. 既 평가 완료된 사례, 특허정보, 재무정보 등에 대한 학습을 기반으로 메타DB를 구축함으로써, 첫째 평가 참조정보로 제공 가능하다는 점, 둘째 정성평가지표 예측을 위한 딥러닝 모형 구축에 적용 가능하다는 점에서 향후 기술평가 시장의 확대와 평가결과의 객관성 제고에 중요한 지능형 시스템으로 발전할 것으로 기대된다.

참고 문헌

- [1] 산업통상자원부-한국산업기술진흥원, *기술금융지원사업 종합성과분석 보고서*, 2018.
- [2] 이준원, 윤점열, “기술력평가모형의 기술금융 활동 적합성 연구,” *기술혁신학회지*, 제20권, 제2호, pp.292-312, 2017.
- [3] 산업통상자원부, *기술평가 실무가이드*, 2021.
- [4] 한국과학기술정보연구원(KISTI), *기술가치평가시스템*, <http://www.starvalue.or.kr>
- [5] 국경완, *인공지능 기술 및 산업 분야별 적용 사례*, 정보통신기획평가원, 2019.
- [6] 정영임, *인공지능(AI) 부활의 동인과 국내의 기술개발 동향*, 정보통신기획평가원, 2016.
- [7] 성태용, 전승표, 김상국, 박현우, “웹기반 지능형 기술가치평가 시스템에 관한 연구,” *지능정보연구*, 제23권, 제1호, pp.23-46, 2017.
- [8] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences(2nd Ed.)*, Lawrence Erlbaum Associates, Inc., 1988.
- [9] K. Harsurinder, “A Systematic Review on Imbalanced Data Challenges in Machine Learning: Applications and Solutions,” *ACM Computing Surveys*, Vol.52, No.4, Article 79, 2019.
- [10] M. I. Jordan and T. M. Mitchell, “Machine learning: Trends, perspectives, and prospects,” *Science*, Vol.349, No.6245, pp.255-260, 2015.
- [11] V. C. Nitesh, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: Synthetic minority

over-sampling technique,” J. Artific. Intell. Res., Vol.16, pp.321-357, 2002.

- [12] M. Galar, A. Fernández, E. Barrenechea, and F. Herrera, “EUSBoost: Enhancing ensembles for highly imbalance data-sets by evolutionary undersampling,” Pattern Recogn., Vol.46, No.12, pp.3460-3471, 2013.
- [13] F. Zhou, S. Yang, H. Fujita, D. Chen, and C. Wen, “Deep learning fault diagnosis method based on global optimization GAN for unbalanced data,” Knowledge-Based Systems, Vol.187, 104837, 2020.
- [14] W. Jiang, Y. Hong, B. Zhou, X. He, and C. Cheng, “A GAN-Based Anomaly Detection Approach for Imbalanced Industrial Time Series,” IEEE Access, Vol.7, pp.143608-143619, 2019.
- [15] E. J. Han and S. Y. Sohn, “Patent valuation based on text mining and survival analysis,” J. Technol. Transf., Vol.40, pp.821-839, 2015.
- [16] 박현우, 이종택, “초기단계 기술의 가치평가 방법론 적용 프레임워크,” 기술혁신학회지, 제15권, 제2호, pp.242-261, 2012.
- [17] C. Giri, S. Thomassey, J. Balkow, and X. Zeng, “Forecasting New Apparel Sales Using Deep Learning and Nonlinear Neural Network Regression,” 2019 International Conference on Engineering, Science, and Industrial Applications (ICESI), pp.1-6, 2019.

저 자 소 개

성 태 응(Tae-Eung Sung)

종신회원



- 2002년 2월 : 서울대학교 전기공학부(공학사)
- 2004년 5월 : (美) 텍사스오스틴 주립대학교 전기컴퓨터공학과(공학석사)
- 2010년 1월 : (美) 코넬대학교 전기컴퓨터공학과(공학박사)
- 2018년 3월 ~ 현재 : 연세대학교(미래) 소프트웨어학부

부교수

〈관심분야〉 : 빅데이터분석, 머신러닝/딥러닝, 기술가치평가, 지능형정보시스템

김 민 승(Min-Seung Kim)

준회원



- 2021년 2월 : 연세대학교(미래) 컴퓨터공학과(공학사)
- 2021년 3월 ~ 현재 : 연세대학교 일반대학원 전산학과(석사과정)

〈관심분야〉 : 기술가치평가, 머신러닝/딥러닝, 금융계량분석, 산업시장분석

이 찬 호(Chan-Ho Lee)

준회원



- 2021년 2월 : 연세대학교(미래) 컴퓨터공학과(공학사)
- 2021년 3월 ~ 현재 : 연세대학교 일반대학원 전산학과(석사과정)

〈관심분야〉 : 기술가치평가, 금융데이터마이닝, 머신러닝/딥러닝, 기업재무분석

최 지 혜(Ji-Hye Choi)

준회원



- 2021년 8월 : 연세대학교(미래) 컴퓨터공학과(공학사)
- 2021년 9월 ~ 현재 : 연세대학교 일반대학원 전산학과(석사과정)

〈관심분야〉 : 머신러닝/딥러닝, 자연어처리, 특허분석

장 용 주(Yong-Ju Jang)

준회원



- 2021년 2월 : 연세대학교(미래) 컴퓨터공학과(공학사)
- 2021년 3월 ~ 현재 : 연세대학교 일반대학원 전산학과(석사과정)

〈관심분야〉 : 기술가치평가, 의료데이터 분석, 머신러닝/딥러닝

이 정 희(Jeong-Hee Lee)

정회원



- 1989년 2월 : 경북대학교 공업화학과(공학사)
- 2018년 8월 : 충북대학교 의생명과학경영융합학과(의생명과학경영융합학 석사)
- 2019년 3월 ~ 현재 : 연세대학교 일반대학원 전산학과(박사과정)

〈관심분야〉 : Business Intelligence, Data mining, Cheminformatics, Preclinical/Clinical Information, Deep Learning, Predictive Maintenance, Computer-Aided Drug Design