

Development of an unsupervised learning-based ESG evaluation process for Korean public institutions without label annotation

Do Hyeok Yoo*, SuJin Bak*

*Assistant Researcher, Advanced Institute of Convergence Technology, Gyeonggi-do, Korea

*Senior Researcher, Advanced Institute of Convergence Technology, Gyeonggi-do, Korea

[Abstract]

This study proposes an unsupervised learning-based clustering model to estimate the ESG ratings of domestic public institutions. To achieve this, the optimal number of clusters was determined by comparing spectral clustering and k-means clustering. These results are guaranteed by calculating the Davies-Bouldin Index (DBI), a model performance index. The DBI values were 0.734 for spectral clustering and 1.715 for k-means clustering, indicating lower values showed better performance. Thus, the superiority of spectral clustering was confirmed. Furthermore, T-test and ANOVA were used to reveal statistically significant differences between ESG non-financial data, and correlation coefficients were used to confirm the relationships between ESG indicators. Based on these results, this study suggests the possibility of estimating the ESG performance ranking of each public institution without existing ESG ratings. This is achieved by calculating the optimal number of clusters, and then determining the sum of averages of the ESG data within each cluster. Therefore, the proposed model can be employed to evaluate the ESG ratings of various domestic public institutions, and it is expected to be useful in domestic sustainable management practice and performance management.

▶ **Key words:** ESG evaluation, Korean Public Institutions, Statistical Analysis, Unsupervised Learning

[요 약]

본 연구는 ESG 등급이 제공되지 않는 국내 공공기관의 ESG 등급을 추정하는 비지도 학습 기반 군집모형을 제안한다. 이를 위해, 스펙트럼 군집과 k-means 군집에서 최적의 클러스터 수를 비교했고, 그 결과의 신뢰성을 보장하기 위해 성능지표인 Davies-Bouldin Index (DBI)를 계산했다. 결과적으로, 스펙트럼 군집과 k-means 군집에서 각각 0.734 및 1.715의 DBI 값을 산출했는데, 이는 값이 작을수록 우수한 성능을 의미하므로 스펙트럼 군집의 우수성을 확인하였다. 게다가, T-검정 및 ANOVA를 이용하여 ESG 비재무 데이터 간 통계적으로 유의미한 차이를 밝혀내고, 상관계수를 이용하여 ESG 항목 간 상관관계를 확인했다. 본 연구는 이러한 결과를 바탕으로 기존 ESG 등급 없이 공공기관별 ESG 성과 순위를 추정할 가능성을 제시한다. 이는 최적의 클러스터 수를 계산한 다음, 각 클러스터 내 ESG 데이터의 평균 총합을 결정함으로써 달성된다. 따라서, 제안된 모델은 다양한 국내 공공기관의 ESG 등급을 평가하는 근거로 활용될 수 있고, 국내 지속가능경영 실천과 성과관리에 유용할 것으로 기대된다.

▶ **주제어:** 환경 사회 지배구조 평가, 공공기관, 통계분석, 비지도 학습

- First Author: Do Hyeok Yoo, Corresponding Author: SuJin Bak
- Do Hyeok Yoo (ydh7019@snu.ac.kr), Advanced Institute of Convergence Technology
- SuJin Bak (soojin7897@snu.ac.kr), Advanced Institute of Convergence Technology
- Received: 2024. 04. 26, Revised: 2024. 05. 14, Accepted: 2024. 05. 17.

I. Introduction

최근 기업의 가치는 재무적 요소뿐 아니라 비재무적 요소의 중요성도 함께 증가하고 있다 [1]. 특히, 기업의 지속 가능한 경영을 위한 핵심 요소인 환경 (Environment), 사회 (Social), 지배구조 (Governance)의 중요성이 부각되면서 국내기업은 ESG 성과를 향상시키기 위해 자발적으로 노력하고 있으며, 이를 등급화하여 체계적으로 관리하고 있다 [2].

국내기업의 ESG 성과는 기업투자와 밀접한 연관성을 맺으면서 기업가치의 향상으로 이어지는 연구가 활발하다 [3-4]. 이러한 현상은 기업의 ESG 성과가 투자자의 관심과 의사결정에 미치는 비중이 커지면서 ESG 성과지표 및 평가등급을 제공하는 평가기관의 투명성과 신뢰성 또한 주목하고 있다. 그러나 평가기관의 ESG 등급을 결정하는데 있어 주요 문제들이 야기되어 왔는데, 이는 평가 기준의 질적 데이터 특성, 기관별 평가 시스템의 가변성, 일반 투자자에 대한 세부 정보, 및 결과의 제한성 등이 존재한다 [5]. 게다가, ESG 경영평가를 필수적으로 수행하는 정책이 강조되면서, 민간기업뿐 아니라 국내 공공기관까지 ESG 성과지표 비중이 점점 확대되고 있다 [6].

최근에는 ESG 경영평가에 인공지능 (Artificial Intelligence; AI) 기술을 도입함으로써 기계학습 (Machine Learning) 및 딥러닝 (Deep Learning) 기반 ESG 경영평가 알고리즘에 관한 연구가 점점 증가하고 있다. Kim et al.(2023)에 따르면 [7], 기업의 재무 정보를 활용하여 심층신경망 (Multi Layer Perceptron; MLP)을 설계하여 ESG 평가를 예측 및 분류하는 모델을 제안했고, 이는 B+ 등급 전후의 ESG 평가 결과를 기준으로 분류 정확도 86%를 달성했다.

또 다른 연구는 재무 데이터 및 ESG 등급에 대한 자료를 확보한 후, 회귀분석 (Regression analysis)을 사용하여 미래 기업가치를 예측하는 모델을 제안하였다 [8]. 연구자는 CatBoost, Bayesian Ridge, Ridge Regression, Linear Regression, 및 Gradient Boosting로 구성된 5개 모델로 실증 분석을 진행하였고, ESG 등급 정보를 입력변수에 포함 여부에 따라 R-Squared, MSE, RMSE, MSLE, 및 MAPE 등의 평가지표를 통해 비교하였다. 결과적으로, ESG 등급 정보가 포함된 모델이 상대적으로 우수한 성능을 보였고, Bayesian Ridge와 Ridge Regression 모델에서 예측 정확도가 가장 높았다.

Krappel, T. et al.(2021)에 따르면 [9], 민간기업의 재무 및 비재무 데이터로 Feedforward, CatBoost, 및

XGBoost로 구성된 3개 모델을 결합한 회귀모형 기반의 앙상블 모델을 개발하였다. 이 모델은 ESG 등급을 예측했는데, 환경·사회·지배구조별 각각 57.5%, 48.7%, 및 23.7%의 예측력을 보였지만, 소규모 데이터로 인해 전반적인 예측의 성능이 낮은 것을 확인할 수 있었다.

이러한 사전 연구는 일반적으로 지도학습 (Supervised Learning) 기반의 분류 또는 회귀 방법론을 적용한 연구가 대다수이고, ESG 등급에 대한 참값 (라벨) 또한 사전정의가 된 상태였다. 게다가, 충분히 학습하기 위해 대용량 데이터가 필수적이지만, 현실 세계에서 ESG 관련 데이터 수는 여전히 부족하고, 라벨링은 되어있지 않는 실정이다.

따라서, 이러한 문제를 해결하기 위해, 본 연구는 데이터 패턴의 경향성을 탐색해 비슷한 성질의 데이터 간 군집화를 이루는 비지도 학습 모델을 제안한다. 국내 공공기관 ESG 운영 및 평가가 점차 주목받고 있는 가운데, 본 연구에서는 AI 비지도 학습 (Unsupervised Learning) 기반 군집분석 알고리즘과 통계분석 기법을 이용하여 공공기관의 ESG 등급 평가 측정을 위한 새로운 방안을 제안하고자 한다. 특히, ESG 등급에 대한 라벨이 제공되지 않는 공공기관 ESG 등급을 추정하는 알고리즘을 개발하여 ESG 경영 전략을 실천하고, 그들의 성과를 효과적으로 관리할 수 있을 것으로 기대한다.

본 논문은 2장에서 ESG 성과와 기업의 (비)재무성과 사이의 상관관계를 다룬 선행연구를 소개한다. 3장은 연구 전반의 단계별 프로세스를 설명하고 시뮬레이션 방법론에 대해 설명하며, 4장에서는 본 시뮬레이션 결과와 참값 (라벨) 없이 ESG 비재무 데이터 평가를 위한 ESG 등급 추정 알고리즘 모델을 제안한다. 마지막으로, 5장에서 본 연구의 시사성 및 사전 연구와의 차별성을 논의하고, 한계 및 향후 연구 방안을 포함하여 결론을 맺는다.

II. Related Works

기업 내 ESG 정책과 전략이 기업의 사회 활동과 성과에 핵심적인 영향을 미치는 요소라는 점은 학계에서 일반적으로 알려진 사실이다. ESG 성과와 기업의 (비)재무성과, 그리고 기업가치 간의 관계에 대한 이론적 기반을 제시한 연구들이 다양한 측면으로 현재 논의되고 있다.

Park, and Sin et al.(2021) 에 따르면 [10], 회귀분석을 이용하여 대기업, 중견기업, 중소기업과 같은 기업의 규모별로 ESG 성과, 재무 상태 및 장단기 기업가치 간의 관계를 분석하였다. 이는 기업 규모와 상관없이, ESG 성

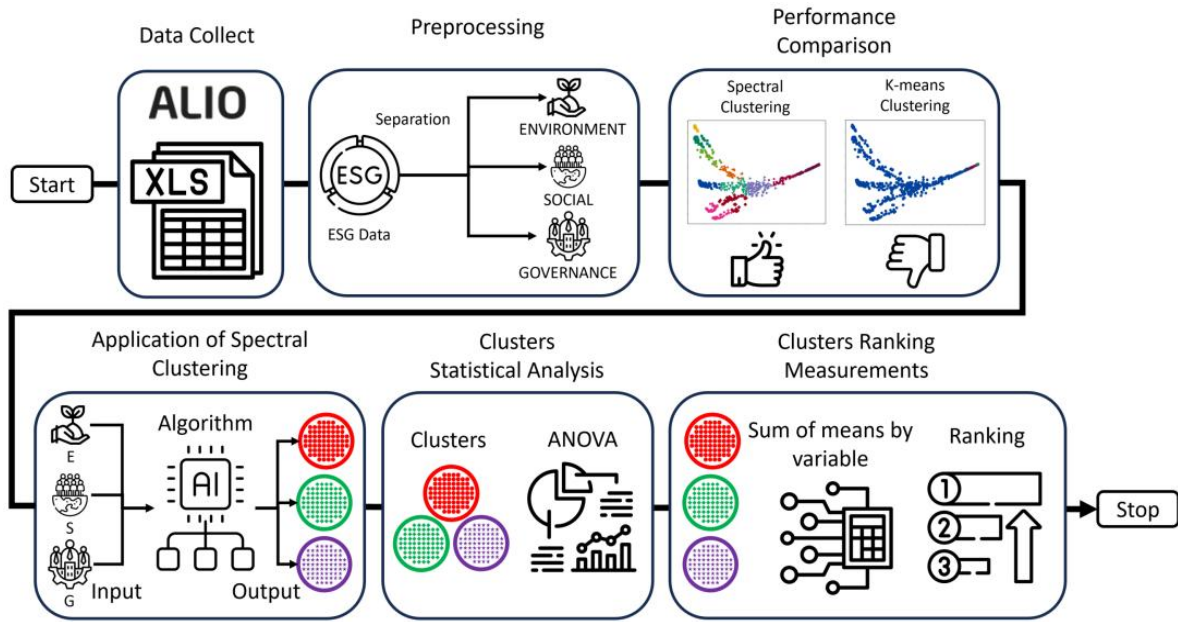


Fig. 1. ESG evaluation process for Korean public institutions

과 및 등급과 기업가치 간 통계적으로 유의미한 차이가 있음을 확인하였다.

유사하게, 최근 연구에 따르면 [11], ESG 경영을 수행하는 기업의 환경등급이 기업의 재무성과에 긍정적인 영향을 미치고 있음을 보여주는데, 이는 ESG 등급이 기업의 자기자본이익률 (Return on Equity; ROE)에 양 (Positive)의 효과를 나타내고 있음이 통계적으로 검증되었다. 또한, 총자산회전율, 부채비율, 매출액증가율, ESG 환경등급 순서로 기업의 ROE에 영향을 미치는 것으로 분석되었다. 즉, 기업의 ESG 성과가 높을수록 기업의 재무 성과도 함께 증대되는 것을 시사한다.

반면에, Kim et al.(2020)에 따르면 [12], ESG 비재무 정보가 재무성과와 기업가치에 아무런 영향이 없음을 증명한 연구사례를 소개했는데, 이는 환경, 사회, 지배구조에 해당하는 ESG 성과와 기업의 총자산에서 당기 순이익이 차지하는 비율을 의미하는 총자산이익률 (Return on Assets; ROA) 간 피어슨 상관관계를 분석한 결과, ESG 공시지수는 모두 종속변수인 ROA에 대해 유의하지 않음을 보여준다. 이는 비재무 정보 공시가 기업의 수익성이나 가치에 영향을 주지 못하는 것을 의미한다.

이러한 기업 내 ESG 성과 및 기업의 재무 활동 간 연관성에 관한 연구들은 항상 모순적인 결과를 동반한다. 이 연구들은 주로 기 등록된 ESG 등급을 이용하여 데이터 분류 및 예측을 위주로 수행된다. 게다가, 국내 공공기관의 ESG 성과와 기업가치 간 상관관계를 분석한 연구 결과는 거의 없으므로, 유관 분야의 발전을 위한 노력이 시급하다.

III. Methods

3.1 Overall ESG Evaluation Process

전체적인 공공기관 ESG 평가 프로세스는 Fig. 1과 같다. 먼저, ESG 비재무 데이터 확보를 위해, 국내 공공기관 경영정보 공개시스템 (All Public Information In-One; ALIO)으로부터 데이터를 확보한 다음, 전처리 과정을 수행한다. 따라서, ALIO 가이드라인에 따라, ESG 비재무 데이터는 크게 3개 항목으로 환경(E), 사회(S), 지배구조(G)로 나뉘고, Table 1과 같이, 국내 179개 공공기관에서 사용하는 소분류 된 ESG 세부 항목으로 분류하였다. 현 공공기관에서 사용되는 비재무 데이터는 ESG 등급에 대한 참값 (라벨)이 없으므로 본 연구에서는 비지도 학습을 바탕으로, 특징 공간에서 군집이 잘 분리되는 것으로 알려진 스펙트럼 군집과 보편적으로 많이 사용되는 k-means 알고리즘을 설계하고 군집 분류 성능을 비교하였다. 그런 다음, 환경, 사회, 지배구조에서 만들어진 평균 군집의 통계적 차이를 비교하기 위해, T-검정 및 ANOVA 통계분석 기법을 활용하였다. 궁극적으로, 세부 항목 평균의 총합을 사용하여 군집에 대한 ESG 순위 측정을 진행하여 공공기관 내 ESG 등급 평가를 추정하는 비지도 학습 모델을 제안한다. 이를 위해, 본 시뮬레이션에서 사용된 컴퓨터의 성능 및 환경은 Window 11 Home 운영체제를 포함 Intel i5-13500KF CPU, 64GB Memory 규격과 같다.

3.2 ESG Datasets and Preprocessing

본 연구는 ALIO에 가입된 179개 공공기관 대상으로 2019년부터 2022년까지의 ESG 비재무 데이터를 분석에 사용하였다. 이러한 ESG 비재무 데이터는 크게 환경 (E), 사회 (S), 지배구조 (G) 항목으로 나뉘고, 세부 항목으로 「온실가스 감축실적」, 「사회공헌활동」, 「중소기업 생 산품 구매실적」 등 기업의 ESG 운영 요소를 판단할 수 있는 77개 항목을 포함하지만, 실제로 77개 ESG 세부 항목을 포함하는 공공기관은 거의 없고, 심지어 ESG 운영을 전혀 하지 않는 공공기관들도 다수 발견되었다.

따라서, 실험 시뮬레이션에 활용하기 위해, 77개의 ESG 세부 항목 중 179개 공공기관에서 공통적으로 확보 가능한 13개 세부 항목에 대한 데이터를 추출하고, 나머지 데이터는 제외시켰다. 이러한 13개 항목은 현 공공기관 내 ESG 지표로써, 확보된 데이터는 Python (Version 3.10.11) language의 Pandas (Version 2.1.1) 및 Numpy (Version 1.24.3) 라이브러리를 활용하여 전처리를 수행하였다. melt() 함수를 이용하여 기관명, 기관유형, 주무기관, 부설기관 여부와 같은 공공기관 데이터를 컬럼 정렬 및 삽입 후, 각 항목의 생성된 컬럼에 사업연도와 매칭시켰다. 동 과정을 모든 공공기관의 ESG 항목별 데이터 추출에 반복적으로 수행한 다음, merge() 함수를 통해 모든 항목을 병합하여 전체 699개의 데이터를 추출하였다. Table 1에 항목-세부 항목을 보여주는데, 이들의 값이 크면 클수록 ESG 지표를 충분히 채웠음을 나타낸다.

3.3 Spectral Clustering

본 연구에서 사용된 스펙트럼 알고리즘은 세 단계로 구분한다. 첫째, 가우시안 커널을 이용하여 유사도(Affinity)를 계산하고 [13], 그 수식은 다음과 같다.

$$k_{\sigma}(x^{(i)}, x^{(j)}) = \exp\left(-\frac{\|x^{(i)} - x^{(j)}\|^2}{2\sigma^2}\right) \quad (1)$$

$$= \exp\left(-\frac{\sum_{k=1}^n (x_k^{(i)} - x_k^{(j)})^2}{2\sigma^2}\right)$$

여기서, $x^{(i)}$ 및 $x^{(j)}$ 는 두 데이터의 포인트이며, $\|x^{(i)} - x^{(j)}\|$ 는 두 데이터 사이의 유클리드 거리이다. 이를 통해 구해진 유사도 행렬을 라플라시안 행렬(graph laplacian)로 수식 (2)와 같이 계산하고 구성한다.

$$D = \text{diag}\left(\sum_{j=1}^n S_{1j}, \dots, \sum_{j=1}^n S_{nj}\right) \quad (2)$$

그 후, 유사도 행렬의 각 행의 합을 대각요소로 갖는 정도 행렬(Degree matrix)로 정의하고, 각 값은 꼭짓점의 차수(Degree) D 를 나타낸다.

둘째, 라플라시안 행렬에 고유값 분해를 적용한다. 목표 군집의 개수를 k 라고 할 때, 분해 시 k 개의 작은 고유값 (Eigen value)에 대응되는 고유벡터 (Eigen vector)를 구한다. 선택된 고유 벡터들을 사용하여 데이터 포인트들을 저차원 공간으로 맵핑한다.

마지막으로, 분해 결과에 k -평균을 적용함으로써 군집화를 진행한다. 이러한 과정을 거쳐 코드화하기 위해 scikit-learn (Version 1.3.1) 라이브러리에서 SpectralEmbedding() 함수를 통해 데이터를 2차원으로 축소 후, KMeans() 함수를 통해 군집화를 진행하여 matplotlib() 함수로 시각화 결과를 산출했다.

Table 1. ESG non-financial data used in this study

Main category	No	Sub-category
Environment (E)	1	Green product purchase amount
	2	Total purchase amount of green and non-green product purchases
	3	Greenhouse gas reduction rate
Social (S)	4	Flexible working status
	5	Number of employees with staggered commuting schedules
	6	Number of female workers
	7	Reduction of working hours during pregnancy scheme
	8	Number of maternity leave users
	9	Number of regular employees
	10	Number of employees with disabilities
Governance (G)	11	Purchasing ratio of products from small enterprises
	12	Purchases amount from small enterprises
	13	Total purchase amount of products and service from small enterprises

3.4 k-means Clustering

스펙트럼 군집화의 성능과 비교하기 위해, k-means 알고리즘을 사용하여 환경, 사회, 지배구조 내 세부 항목 간 클러스터링을 진행하였고, 또한 통합 ESG에 대한 클러스터링을 수행하였다. 이 방법은 각각의 관측치 벡터 (Observation vector)를 가장 가까운 중심 (또는 평균)을 지닌 군집 (Cluster)에 할당하는 알고리즘으로써, 다음과 같은 수식으로 표현된다 [14].

$$W(S, C) = \sum_{k=1}^K \sum_{i \in S_k} \|y_i - c_k\|^2 \quad (1)$$

여기서, S 는 다차원 특징 공간에서 $y_i(i \in I)$ 로 표현되는 개체 집합의 k 개 클러스터 분할을 의미한다. 각 차원은 비어있거나 중첩되지 않은 S_k 개의 군집으로 구성되고, 이때 S_k 군집의 중심은 $c_k(k=1, \dots, K)$ 이다. 이 수식을 컴퓨팅 과정에 적용하기 위해, scikit-learn 라이브러리에서 제공하는 KMeans() 함수를 사용하여 군집화하고, matplotlib() 함수로 시각화 결과를 도출하였다.

3.5 Clustering model verification analysis

본 연구에서는 계산된 두 가지 클러스터링 모델은 군집 간 데이터의 분리도(Cohesion)와 응집도(Separation)를 동시에 고려한 Davies-Bouldin Index (DBI) 지표를 바탕으로 군집화 성능을 평가하였다. 여기서 군집의 분리도는 각 군집 간 이질성을 최대화한 정도를 나타내고, 응집도는 동일 군집 내 유사성을 최대화한 정도를 의미한다. 본 연구는 Von Luxburg et al.(2007)에 따라 [13], DBI를 계산하였고, 이는 고유값 간격 (Eigen-Gap)을 계산하여 최적의 k 값을 구할 수 있다. 따라서, k 의 값이 작을수록 군집화의 성능이 뛰어나다. DBI의 계산식은 다음과 같다.

$$DBI = \frac{1}{k} \sum_{i=1}^k \max \left(\frac{\Delta(X_i) + \Delta(X_j)}{\delta(X_i, X_j)} \right) \quad (1)$$

여기서 $\delta(X_i, X_j)$ 는 클러스터 간 거리로써 클러스터 X_i 와 X_j 사이의 거리 분리도를 나타내며, ΔX_k 는 클러스터 X_k 내의 클러스터 내 거리 응집도이다. 이를 scikit-learn라이브러리의 davues_bouldin_score() 함수로 계산했다.

3.6 Statistical Analysis Using T-test and ANOVA

본 연구에서 수집 및 전처리된 자료는 SPSS 26 Program을 이용하여 분석하였다. 전 과정은 Oh et al.(2023)의 방법론에 따라 [15], 연구변수인 환경, 사회, 지배구조, 및 세 개 항목을 포함한 통합 ESG 항목별 세부 항목의 군집 간 차이를 분석하기 위해, 빈도분석과 기술통계 분석을 수행하였다. 군집 유형별 데이터 집단 간 비교를 위해 데이터 특성을 파악하고자 Levene 등분산 검정을 수행하였다. 그런 다음, 군집유형에 따른 ESG 항목 변수의 평균값 비교를 위해 T-검정과 Welch 일원분산분석 (Welch One-way ANOVA)을 실시하였고, 집단 간 차이 비교를 위해 사후분석 (Games-Howell)을 실시하였다. 추가적으로, ESG 등급 라벨이 없는 비재무 데이터를 사용하

여 비지도 학습 기반의 군집화 모델을 개발하기 위해, 환경·사회·지배구조 항목 간 요인별 피어슨 상관관계 (Pearson's Correlation) 분석을 수행하였다.

IV. Results

4.1 Clustering model verification

본 연구에 사용된 두 가지 군집모델에서 최적의 군집 수를 계산한 결과는 Fig. 2와 같다. 스펙트럼 군집은 최적의 k 를 구하는데 고유값 간격을 사용하였으며, 이는 (a) 환경, (b) 사회, 및 (c) 지배구조, (d) 통합 ESG 순으로 각각 2개, 13개, 5개, 및 17개로 계산되었다.

마찬가지로, K-means 군집의 최적화된 k 를 도출하기 위해, DBI를 이용하여 (d) 환경, (e) 사회, (f) 지배구조, (g) 통합 ESG 순으로 4개, 2개, 2개, 및 11개를 각각 계산했다. 여기서 모든 그래프의 가로축은 최적의 k 개수를 의미하고, 스펙트럼 군집에서 세로축은 고유값 간격을 의미한다. K-means 군집의 경우, 세로축은 DBI 값을 나타낸다. Table 2는 구체적인 DBI 값을 보여준다. 스펙트럼 군집에서 DBI는 환경, 사회, 지배구조, 및 통합 ESG 순으로 0.764, 0.530, 0.574, 및 0.734의 값을 각각 산출했다. K-means 군집의 경우, 0.621, 0.596, 2.717, 및 1.715로 계산되었다. DBI는 0에 가까울수록 우수한 성능임을 의미하는데, 스펙트럼 군집에서 사회(S) 항목의 군집화가 우수함을 나타낸다. 유사하게, K-means 군집에서도 사회(S) 항목의 군집화가 가장 뛰어난 것을 알 수 있다. 이러한 결과는 사회(S) 항목의 변인 수가 가장 많기 때문일 것으로 추측된다. 또한, 본 연구에서 최종적으로 통합 ESG에서 K-means 모델과 비교하여 스펙트럼 군집의 DBI 성능이 뛰어난을 확인하였다.

Table 2. Clustering performance by DBI values

Items	Spectral Clustering	K-means Clustering
	DB index (DBI)	
E	0.764	0.621
S	0.530	0.596
G	0.574	2.717
ESG	0.734	1.715

4.2 Clustering model visualization results

Fig 3은 두 개 비지도 학습 모델의 군집화 결과에 대한 matplotlib() 함수를 사용한 시각화 결과를 보여준다. (a)-(d)는 스펙트럼 군집의 환경, 사회, 지배구조, 및 통합

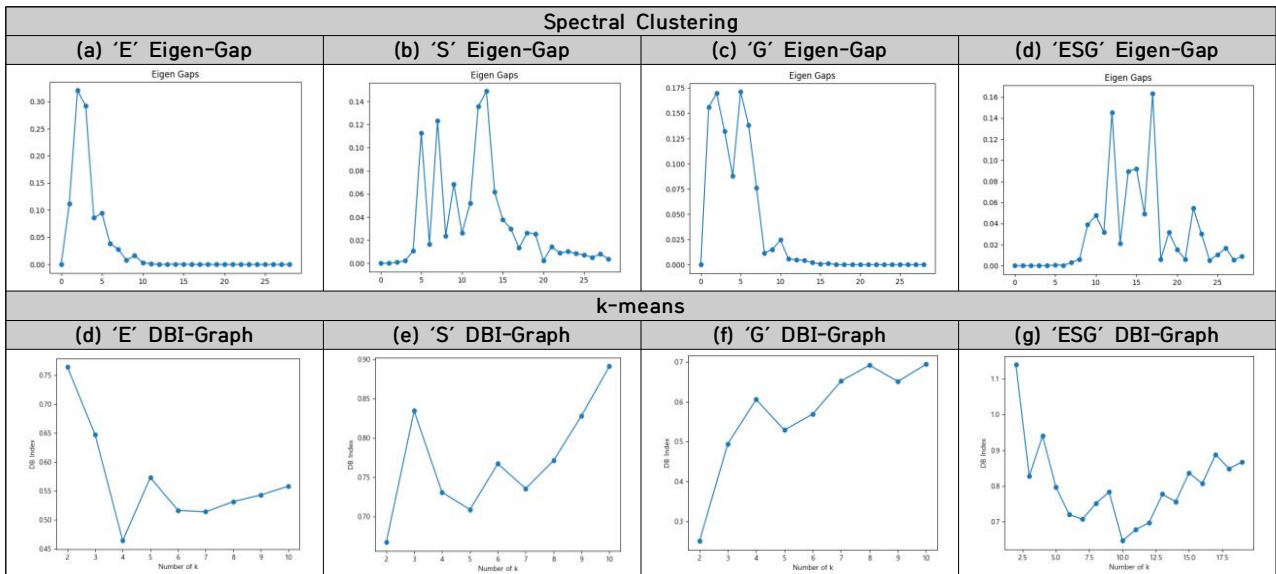


Fig. 2. Graphs of optimal k-value derivation by clustering algorithm

ESG의 가시화 결과를 각각 나타내고, (e)-(h)는 k-means 군집의 환경, 사회, 지배구조, 및 통합 ESG의 시각화 결과를 보여준다. 각 그래프 내 레전드는 항목별 최적의 k 개수를 0에서 $k-1$ 개수로 표현했다.

결과적으로, Fig. 3의 (a)-(d)의 모든 수치는 각 항목 내 군집 간 겹치는 데이터 포인트가 거의 없으므로 스펙트럼 군집화가 적절히 이루어진 것을 알 수 있는 반면, (e)-(h)는 데이터 포인트가 겹치는 부분이 발생하여 k-means 군집화 과정에서 문제가 있음을 확인하였는데, 이는 k-means 군집의 비재무 데이터 간 편향이 생겨 불균형 문제가 발생함을 알 수 있다. 이를 고려하여, 상대적으로 스펙트럼 군집 알고리즘 우수성을 확인하였다.

4.3 Statistical results

본 연구는 항목 내 요인별 통계적 차이를 분석하기 위해, 빈도분석 및 기술통계 분석을 수행하였다. 그 결과, 각 항목에서 결측값 없이, 699개 유효한 데이터를 계산했고, 유효 퍼센트의 총합은 100%가 되는데, 이는 항목별 데이터 수의 적절성을 보여준다.

Table 3은 세부 항목 간 통계적 차이를 비교하였다. 여기서 L 은 Levene 통계량을 나타내고, t 는 t 통계량을 보여주고, p 는 유의확률을 의미한다. 환경에서는 「온실가스 감축률 (No. 3)」에서만 Levene 등분산을 가정하였고, 나머지 변인에 대해서는 Levene 등분산을 가정하지 않는 상태에서 T-검정을 수행하였다. 환경 항목에서 평균 「녹색

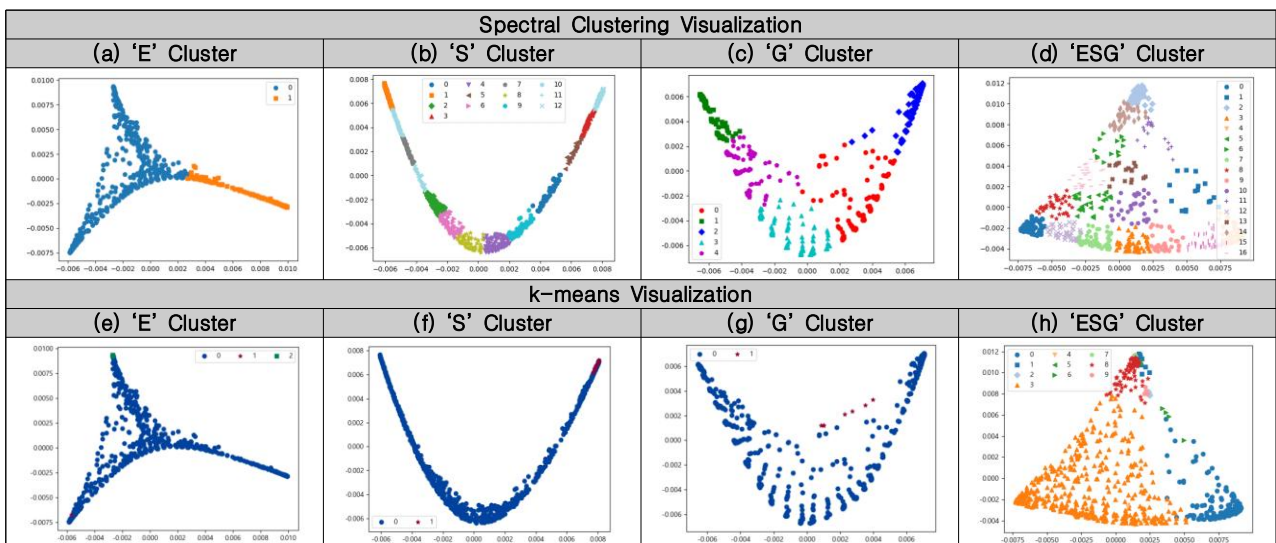


Fig. 3. Visualization results by spectral clustering and k-means

Table 3. 'E', 'S', and 'G' Statistical analysis results

	Values	No. 1		No. 2		No. 3		
	E	L	20.966		21.515		.438	
	p	.000		.000		.508		
	t	5.872		6.017		-5.173		
	p	.000		.000		.000		
S	Values	No. 4	No. 5	No. 6	No. 7	No. 8	No. 9	No. 10
	L	40.961	43.253	102.156	53.231	85.467	43.050	32.513
	p	.000	.000	.000	.000	.000	.000	.000
	F	116.352	85.217	54.689	36.980	59.773	100.405	66.054
	p	.000	.000	.000	.000	.000	.000	.000
G	Values	No. 11		No. 12		No. 13		
	L	16.935		14.247		162.298		
	p	.000		.000		.000		
	F	11.912		11.844		1259.204		
	p	.000		.000		.000		

제품구매액 (No. 1)», 「녹색제품구매실적총구매액 (No. 2)», 및 「온실가스감축률 (No. 3)»의 세부 항목에서 공공기관별 유의미한 차이가 있음을 발견하였다. 유사하게, 사회, 지배구조 항목에서 Levene 통계량과 등분산을 가정하지 않는 상태에서 Welch 일원분산분석 (Welch One-way ANOVA)을 실시한 결과, 공공기관에 따른 세부 항목별 유의미한 평균 차이를 보여준다. 다시 말해, 환경 내 일부 변인, 사회, 지배구조에 대한 분산의 동질성 가정을 검정한 결과, 모든 항목에서 유의수준의 결과를 나타내므로 군집 간 평균의 유의미 $0.001 < p^{***}$ 한 차이를 보여주었고, 이는 ESG 항목 내 요인별 공공기관에 따른 평균의 차이가 있음을 시사한다.

유사하게, Table 4는 환경, 사회 지배구조에 대한 통합 ESG의 공공기관별 평균 차이를 통계적으로 보여주는데, 모든 요인에서 통계적으로 유의미한 평균 차이를 보여준다. 본 결과의 신뢰성을 검증하기 위해, 사후검정 (Games-Howell test)을 수행하였고, 모든 군집에서 유의미한 통계적 차이를 보여준다. 즉, 모든 군집 간의 평균에 차이가 있음을 나타내며, 이를 활용하여 군집 간의 순위를 매길 수 있는 가능성을 보여준다. 본 연구 결과를 통해 ESG 정책을 수행하는 공공기관별 ESG 전략의 차이가 있음이 밝혀졌고, 강조하는 ESG 복지정책이 다를 수 있음을 시사한다.

4.4 Correlations between sub-category data

Fig. 4는 환경, 사회, 지배구조 내 요인 간 상관관계를 각각 보여준다. 여기서, 상관계수가 1에 가까울수록 양의 상관관계를 나타내고, -1에 가까울수록 음의 상관관계를 보여준다. 0은 관계없음을 의미한다.

Table 4. 'ESG' Statistical analysis results

Variables	L	F	p<.001***
No. 1	64.350	14.258	.000***
No. 2	66.061	12.153	
No. 3	2.001	2.103	
No. 4	31.400	43.970	
No. 5	33.396	29.853	
No. 6	84.567	14.560	
No. 7	43.592	12.449	
No. 8	71.016	14.076	
No. 9	31.867	39.699	
No. 10	23.952	29.622	
No. 11	26.769	424.046	
No. 12	125.273	14.915	
No. 13	67.152	15.482	

결과적으로, 환경 영역에서 3개 요인별 피어슨 상관계수를 분석해보면, 「녹색제품구매액 (No. 1)»이 클수록 「녹색제품구매실적총구매액 (No. 2)»도 증가하는 양의 상관관계를 나타내는 반면, 「녹색제품구매액»과 「녹색제품구매실적총구매액»은 「온실가스감축률 (No. 3)»과는 0.01 수준의 양의 상관관계를 보이지만, 이는 「온실가스감축률»의 값이 백분율로 나머지 요인에 비해 작으므로 사실상 상관관계가 거의 없음을 의미한다.

사회영역에서는 모든 요인에서 양의 상관관계를 보이는데 이는 하나의 요인에서 값이 증가할수록 다른 모든 요인도 함께 증가하는 양의 상관관계를 보여준다.

마지막으로, 지배구조 영역에서 「중소기업제품구매액 (No. 11)»이 클수록 「중소기업제품구매실적총구매액 (No. 12)»도 증가하지만, 「중소기업제품구매액»과 「중소기업제품구매실적총구매액»이 감소할수록 「중소기업제품구매비율 (No. 13)»도 감소하는 음의 관계를 보여준다. 결국, 피어슨 상관계수 분석의 결과로 요인별 특성에 따른 증감패턴을 알 수 있었고, 이를 통해 비재무 데이터 간 경향성을 분석할 수 있다.

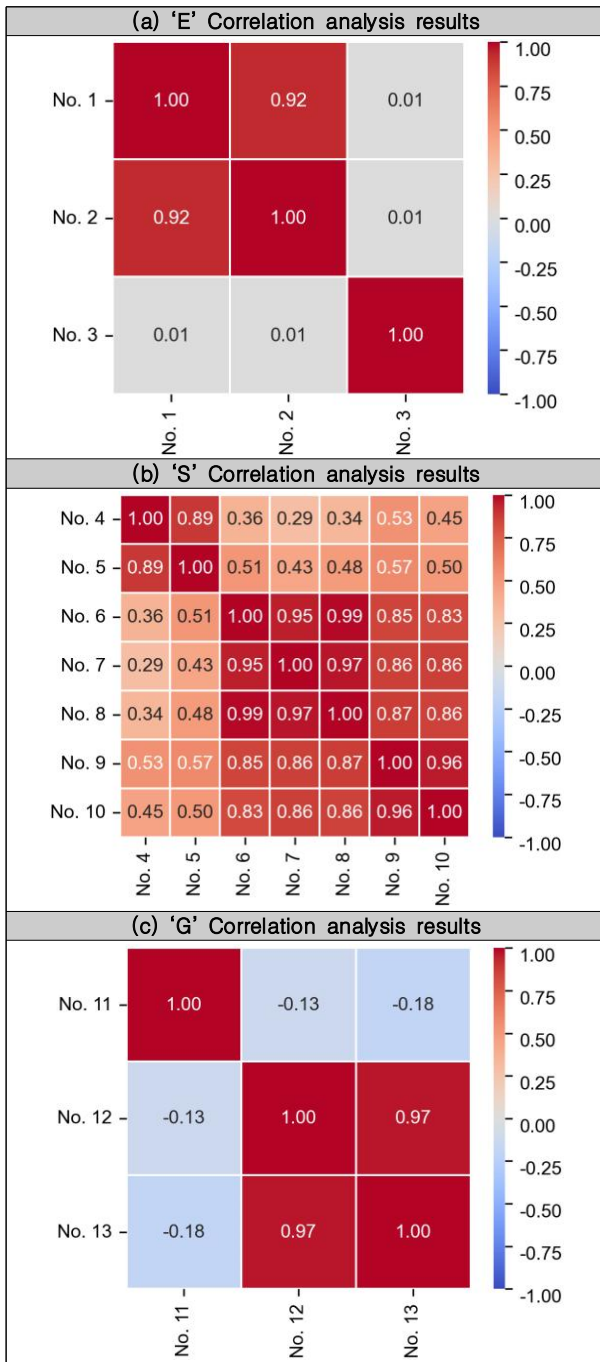


Fig. 4. E-S-G using Correlation coefficients

4.5 Proposed unlabeled ESG rating algorithm based on unsupervised learning

본 연구에서는 공공기관의 ESG 비재무 데이터에 대해 환경, 사회, 지배구조 부문의 ESG 등급을 추정하기 위해 13개 ESG 항목의 데이터 특성을 활용했다. ALIO에 따르면, 13가지 ESG 지표의 값이 클수록 ESG 등급이 높은 것으로 나타나는데, 이러한 추세를 바탕으로 우수한 성능을 보인 스펙트럼 군집을 사용하여 최적의 클러스터 수를 계산했다. 그런 다음, 각 클러스터 내 포함된 모든 기관의 E,

S, G, ESG 항목에 대한 클러스터별 평균의 총합이 계산되었다. 총합이 가장 높은 군집은 A 등급으로, 평균의 총합이 가장 낮은 군집은 최하 등급으로 지정되었다.

Table 5. Cluster rank of 'E', 'S', 'G', and 'ESG'

Items	Number of Cluster	Sum of Average	Estimated Rank
E	0	3988.260	A
	1	396.103	B
S	0	1853.473	D
	1	193.899	M
	2	526.400	I
	3	4245.850	B
	4	1047.813	F
	5	2677.215	C
	6	567.931	H
	7	330.700	K
	8	700.778	G
	9	1468.409	E
	10	252.520	L
	11	391.381	J
12	9266.644	A	
G	0	309220.884	A
	1	40100.236	D
	2	218476.449	B
	3	39100.616	E
ESG	4	127792.954	C
	0	25082.090	P
	1	635361.400	B
	2	215248.700	C
	3	32356.330	M
	4	113117.000	E
	5	44959.080	K
	6	71664.840	I
	7	14141.800	Q
	8	64464.020	J
	9	28059.540	O
	10	101098.600	G
	11	2697605.000	A
	12	32025.540	N
	13	77576.530	H
	14	126338.400	D
15	44166.580	L	
16	112447.000	F	

결과적으로, Table. 5는 환경, 사회, 지배구조별로 최적의 k수 만큼 ESG 항목별 평균의 총합을 보여주고, 해당 군집에 적합한 추정 등급을 나타낸다. 군집 평균 결과의 값이 큰 순서대로 높은 ESG 등급을 부여할 수 있다. 환경, 사회, 지배구조, 및 통합 ESG에서 각각 0, 12, 0, 및 11번 클러스터에서 가장 높은 ESG 등급인 A등급을 받았고, 1, 1, 3, 및 7번 클러스터에서 가장 낮은 등급을 받았다. 이를 통해, 높은 등급과 낮은 등급을 추정할 수 있게 되면서, 라벨 없는 공공기관의 비재무 데이터를 동 방식으로 계산 후, 그 평균의 총합 값을 통해 ESG 순위화가 가능할 것으

로 기대된다. 따라서, ESG 등급의 라벨이 지정되지 않은 공공기관의 ESG 등급 평가를 추정할 수 있다.

V. Conclusions

ESG의 중요성이 부각되면서 ESG 운영 전략은 기관의 가치를 평가하는 중요한 요소 중 하나로 알려져 있다. 그러나 비재무 데이터에 의존하는 현행 공공기관 ESG 평가 시스템은 여전히 미흡하고, 활용 가능한 데이터도 부족한 실정이다. 이러한 문제를 보완하기 위해, 본 연구는 ALIO에서 수집한 공공기관 ESG 비재무 데이터를 사용하여 스펙트럼 군집과 k-means 군집의 DBI 값을 서로 비교함으로써 스펙트럼 군집의 우수성을 확인하였다. 또한, 비재무 데이터 지표 사이 상관계수를 분석하고, T-검정 및 ANOVA를 통한 ESG 데이터 간 증감패턴을 확인하였다. 이러한 결과를 바탕으로 본 연구에서는 ESG 항목 지표에 대한 데이터 구조와 성질을 파악하여 스펙트럼 클러스터링 모델을 중심으로 라벨 없이 ESG 등급을 평가할 수 있는 알고리즘을 제안한다. 본 연구는 구체적으로 다음과 같은 시사점을 제시한다.

첫째, 본 연구는 국내 공공기관을 위한 비지도 학습 기반 ESG 평가 모델을 제안한다. 이 모델은 데이터 샘플과 라벨을 기반으로 학습 과정을 필요로 하는 전통적인 통계 분석이나 심층신경망 모델과 달리, 항목별 클러스터의 평균 ESG 성과를 계산한다. 이러한 접근 방식은 공공기관의 특정 ESG 라벨 없이도 ESG 성과를 순위 매길 수 있는 가능성을 시사하며, ESG 비재무 데이터 내 구조와 패턴을 식별하는 데 있어서 비지도 학습의 강점을 보여준다.

둘째, 공공기관의 ESG 경영 강화를 목표로 하는 K-ESG 가이드라인에 따라, 국내 공공기관의 특성과 자율성을 보장하는 ESG 경영 진단 항목을 강화하고 확장하기 위한 지표를 개발하는 데 의의가 있다. 본 연구 모델은 기관별 ESG 비재무 데이터 항목에 등급을 부여할 수 있어, 추가하거나 삭제해야 할 ESG 항목을 비교 분석할 수 있으며, 이를 통해 적절한 ESG 항목 평가 시스템을 구축하는데 활용될 수 있다는 점에서 의미가 있다.

그러나 본 연구에서 사용된 스펙트럼 군집의 특성상 데이터 특성과 유사성을 결정하는 과정에서 신뢰성 검증 과정이 생략되었다. 즉, 새로운 데이터에 대한 라벨이 없어서 군집 결과의 신뢰성을 절대적으로 평가할 수 없었다. 이러한 한계를 해결하기 위해, 모델의 교차검증을 수행하고 이를 지도학습 기반 분류 문제로 확장하여 신뢰성 검증

평가를 수행할 필요가 있다.

따라서, 향후 연구에서 이러한 문제를 해결함으로써 국내 ESG 평가시장의 투명성과 신뢰성이 크게 향상될 것으로 기대한다.

ACKNOWLEDGEMENT

This work was supported in part by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2023S1A5A8076043) and in part by the Advanced Institute of Convergence Technology (23AI015).

REFERENCES

- [1] Ji-Yun Kim et al., "Development of ESG Index Scale Tool for Sustainable Management at the University." *Proceedings of the Korean Society of Computer Information Conference*, Vol. 30, No. 1, pp. 215-216, Daejeon, South Korea, 2022.
- [2] Hoyjung Kim, and Lee Junseok, "A study on the gap between ESG implementation and evaluation: focusing on the ESG performance indicators." *Institute of Public Policy and Administration*, Vol. 35, No. 4, pp. 199-225, 2021. DOI: 10.17327/ippa.2021.35.4.008
- [3] Won Kang, and Mookwon Jung, "Effect of ESG Activities and Firm's Financial Characteristics." *Korean J Financ Stud*, Vol. 49, No. 5, pp. 681-707, Oct. 2020. DOI: <https://doi.org/10.26845/KJFS.2020.10.49.5.681>
- [4] Oh Sang-hui, "A Study on the Effect of Agency Rating and ESG Rating on Corporate Value." *TAX ACCOUNTING RESEARCH*, Vol. 69, No. 0, pp. 125-144, 2021. DOI: 10.35349/tar.2021.69.006
- [5] Consolandi, Costanza et al., "Material ESG outcomes and SDG externalities: Evaluating the health care sector's contribution to the SDGs." *Organization & Environment* Vol. 33, No. 4, pp. 511-533, 2020.
- [6] Lim, Hyo-Chang, "Exploratory Study on ESG Performance Management of Public Enterprises." *Korean Management Consulting Review*, Vol. 24, No. 1, pp. 15-28, Feb. 2024.
- [7] Min-Seung Kim, Seung-Hwan Moon, Sungwon Choi, "Predicting Future ESG Performance using Past Corporate Financial Information: Application of Deep Neural Networks." *J. Intell. Inf. Syst.*, Vol. 29, No. 2, pp. 85-100, Jun. 2023. DOI: <http://dx.doi.org/10.13088/jiis.2023.29.2.085>

- [8] Kyung Gu, Kang, Cheong Yeul, Park, Hyung Jong, Na, "A Comparative Study of Machine Learning-Based Future Enterprise Value Prediction Models: Impact of ESG Evaluation Rating." *Korean Journal of Business Administration*, Vol. 36, No. 9, pp. 1515-1537, 2023.
- [9] Krappel, Tim, Alex Bogun, and Damian Borth, "Heterogeneous ensemble for ESG ratings prediction." arXiv preprint arXiv:2109.10085, 2021. DOI: <https://doi.org/10.48550/arXiv.2109.10085>
- [10] Soonae Park, and Eunhae Shin, "Analysis of the Relationship between ESG Performance and Corporate Financial Values Focusing on SMEs." *Journal of Environmental Policy and Administration*, Vol. 29, No. 4, pp. 151-199, Dec. 2021. DOI: 10.15301/jepa.2021.29.4.151
- [11] In-Young Cho, "A Study on Corporate ESG Environmental Ratings and Financial Performance." *Logos Management Review*, Vol. 21, No. 2, pp. 269-290, 2023. DOI: 10.22724/LMR.2023.21.2.269
- [12] Yun Kyung Kim, "Effects of Non-financial Information Disclosure on Firm Performance and Firm Value" *Korea Society For Regulatory Studies*, Vol. 29, No. 1, pp. 35-59, Jun. 2020. DOI: <http://dx.doi.org/10.22954/ksrs.2020.29.1.002>
- [13] Von Luxburg, and Ulrike, "A tutorial on spectral clustering." *Statistics and computing* Vol. 17, pp. 395-416, Aug. 2007. DOI: <https://doi.org/10.1007/s11222-007-9033-z>
- [14] T. M. Kodinariya, and P. R. Makwana, "Review on Determining Number of Cluster in K-Means Clustering," *Int. J. Adv. Res. Comput. Sci. Manag. Stud.*, Vol. 1, No. 6, pp. 90-95, Nov. 2013.
- [15] Oh, Hee-Kyun, Kang, Jae-Wan, "Development of a diagnostic index model for regional tourism activation using higher-order confirmatory factor analysis: application of mobile communication and credit card big data." *Journal of Hospitality and Tourism Studies*, Vol. 25, No. 1, pp. 77-94, 2023.

Authors



Do Hyeok Yoo received the Associate degree in Computer Software from Yeonsung University, South Korea, in 2023. He is currently working toward the B.S. degree in Computer Software from Yeonsung

University, South Korea. He joined the Advanced Institute of Convergence Technology at Seoul National University, Gyeonggi-do, South Korea, in 2023. He is currently a Research Engineer in Artificial Intelligence Signal Processing Laboratory (AISPL). He is interested in AI, deep learning, computer vision, and so on.



SuJin Bak received a B.S. degree in computer engineering from Jeonbuk National University, South Korea, in 2018, and subsequently earned Ph.D. degrees in brain and cognitive engineering from Korea University, South Korea, in 2023.

Dr. Bak joined the Advanced Institute of Convergence Technology at Seoul National University, Gyeonggi-do, South Korea, in 2023. She is currently a Director of Artificial Intelligence Signal Processing Laboratory (AISPL). Her primary research interests involve data algorithm designs using machine learning and deep neural network techniques.