

물류공동화 활성화를 위한 빅데이터 마이닝 적용 연구 : AHP 기법을 중심으로

박영현

경남대학교 무역물류학과 조교수

이재호

성균관대학교 무역연구소 연구원

김경우

(사)양산발전연구원 원장

Study on the Application of Big Data Mining to Activate Physical Distribution Cooperation : Focusing AHP Technique

Young-Hyun Pak^a, Jae-Ho Lee^b, Kyeong-Woo Kim^c

^aDepartment of International Trade and logistics, Kyungnam University, South Korea

^bInstitute of Foreign Trade, Sungkyunkwan University, South Korea

^cInstitute of Yangsan Development Research, South Korea

Received 30 September 2021, Revised 26 October 2021, Accepted 28 October 2021

Abstract

The technological development in the era of the 4th industrial revolution is changing the paradigm of various industries. Various technologies such as big data, cloud, artificial intelligence, virtual reality, and the Internet of Things are used, creating synergy effects with existing industries, creating radical development and value creation. Among them, the logistics sector has been greatly influenced by quantitative data from the past and has been continuously accumulating and managing data, so it is highly likely to be linked with big data analysis and has a high utilization effect. The modern advanced technology has developed together with the data mining technology to discover hidden patterns and new correlations in such big data, and through this, meaningful results are being derived. Therefore, data mining occupies an important part in big data analysis, and this study tried to analyze data mining techniques that can contribute to the logistics field and common logistics using these data mining technologies.

Therefore, by using the AHP technique, it was attempted to derive priorities for each type of efficient data mining for logisticalization, and R program and R Studio were used as tools to analyze this. Criteria of AHP method set association analysis, cluster analysis, decision tree method, artificial neural network method, web mining, and opinion mining. For the alternatives, common transport and delivery, common logistics center, common logistics information system, and common logistics partnership were set as factors.

Keywords: Big Data Analytic, SCM, AHP, Physical Distribution Cooperation

JEL Classifications: D81, N70

^a First Author, E-mail: pyh21@kyungnam.ac.kr

^b Corresponding Author: E-mail: skku005@gmail.com

^c Co-Author: E-mail: runner8627@hanmail.net

I. 서론

4차 산업혁명 시대라고 통칭되는 현대 사회는 ‘초연결’, ‘초지능’, ‘초융합’으로 대표되며 빅데이터(Big Data), 클라우드(Cloud), 인공지능(AI), 가상현실(VR), 사물인터넷(IoT) 등의 기술이 주도할 것으로 기대되고 있다. 4차 산업혁명이라는 용어는 2016년 스위스에서 열린 다보스 포럼(Davos Forum)에서 당시 의장이었던 클라우스 슈밥(Klaus Schwab, Germany)이 Industry 4.0이 전 세계 질서를 새롭게 만드는 동인이 될 것이라고 언급하면서 이슈화됐다. 4차 산업혁명의 주요 기술 중에서 빅데이터 분석은 클라우드 컴퓨팅 등과 함께 큰 주목을 받고 있으며, 인공지능의 등장과 함께 그 파급 효과가 더욱 커지고 있다.

빅데이터는 물류·유통분야에서도 많은 영향을 미칠 것으로 여겨지며, 다양한 활용 가능성이 제시되고 있다. 이미 기존의 물류 분야에서 정보 및 정보시스템의 중요성이 강조되어 왔기 때문에 이러한 빅데이터와의 연계는 더 효율적일 것으로 기대되어지고 있다. 빅데이터 분석과 활용을 위한 연구로서 Lee(2016)는 빅데이터 적용에 적합한 산업을 AHP 방식을 통해 탐색하였으며, 빅데이터 활용 분야를 제조·서비스·공공 분야로 구분하였을 때 서비스 분야에서 물류 분야의 빅데이터 활용이 가장 적합한 것으로 평가되었다(Lee Sang-Won, Kim Sung-Hyun, 2016). Kim(2017)은 물류 분야에서 DHL 및 일부 기업들의 빅데이터 활용 사례를 바탕으로 운영의 효율성, 고객경험, 신사업 모델을 요인으로 하는 가치 모델을 제시하였다(Kim, Seung-Wook, 2017). Shin and Kim(2018)은 빅데이터의 국내외 기술동향 및 사례를 분석하고, 국내외의 물류 빅데이터 활용실태를 바탕으로 물류 빅데이터 활성화에 대한 제도적 애로요인 및 개선방안을 제시하였다. 이처럼 학계에서는 물류에서 빅데이터 분석을 활용하기 위한 사례 연구가 활발하게 이루어지고 있다(Shin, Na-la and Kim, Yong-Jin, 2018). Park, Jung, Kwon(2019)은 의료 SCM 경쟁역량 강화를 위해 물류공동화 시스템 도입의 필요성을 강조하였으며, 의료 SCM의 경쟁역량

요인을 협업시스템, 가격리더십, 인도속도, 프로세스유연성으로 구성하여 이에 대한 기대성과가 물류공동화 사용의도로 연결됨을 분석하였다(Park, Kwang-O, Jung, Dae-Hyun, Kwon, Sang-Min, 2019). Lee(2021)는 전자상거래의 발달과 라스트마일의 효율성 증대로 무인서비스가 증가하였으며 이로 인한 스마트 물류의 확대가 요구되어진다고 하였다. 또한 이를 위해 ICT 첨단기술과 물류기술이 융합되어야 하며 물리적 자동화, 빅데이터 분석을 통한 의사결정의 자동화 달성에 대하여 언급하였다(Lee, Eun-Jung, 2021).

다만, 물류 분야에서 빅데이터 연구는 기업 또는 공공기관이 보유하고 있는 기술의 활용 사례 소개 또는 중요성 및 필요성에 대한 연구 위주로 이루어지고 있다. 빅데이터라는 대규모 데이터에서 경제적 가치를 추출하고 분석하는 마이닝 기술에 대한 물류적 접근 및 학술적 연구는 아직 미흡한 실정이다. 일반적으로 빅데이터 분석(Big Data Analytic, BDA)은 대규모 데이터라는 측면에서 관심을 받고 있으나, 4차 산업혁명 시대를 맞이하며 이슈가 되고 있는 빅데이터 분석은 가치 있는 데이터를 추출하고 분석할 수 있는 기술이 개발되고 있다는 점에 더 중요한 의미를 두고 있다. 이러한 기술을 데이터 마이닝(Data Mining)이라 하며 이는 데이터에 숨겨진 패턴과 상관관계를 발견하여 새로운 정보를 발견해 내는 과정을 의미한다.

따라서 물류 분야에 있어서 데이터 마이닝에 대한 연구가 이루어진다면 기존의 물류 분야 및 물류 공동화 과정에 대한 추가적인 이해가 이루어질 수 있을 것으로 여겨진다. 또한 물류 분야 데이터 마이닝 연구는 기존 물류 데이터를 효율적으로 사용할 수 있는 방안에 대한 이해가 이루어질 수 있으며, 보다 정교한 데이터 추출에 활용될 수 있으므로 이 분야에 대한 연구가 요구되어지고 있다.

본 연구에서는 다양한 데이터 마이닝 기법 중에서 물류 분야에서 활용할 수 있는 데이터 마이닝 기법을 도출하고, 활용하기 위한 우선순위를 AHP 기법을 이용하여 분석하고자 하였다. AHP 기법은 전문가의 견해를 반영하는 의사결정 기법으로 분석과정에서 의사결정자의

일관성을 판단할 수 있는 지표를 제공하여 절차가 합리적이고, 데이터 마이닝에 대한 물류 분야 학술적 연구가 부족한 상황에서 합리적으로 활용할 수 있는 통계 기법이기때 이를 활용하였다. 다만 물류 분야 전방위적 연구는 물류의 세부적인 분야별 상이한 차이가 존재할 수 있으며, 데이터 마이닝을 적용하는 구체적인 목표 설정을 위하여 빅데이터 활용이 주요한 요인으로 적용될 수 있는 물류공동화 활성화를 목표로 설정하여 연구를 진행하였다.

II. 빅데이터 마이닝

데이터 마이닝(Data Mining)은 대용량의 데이터로부터 자동 또는 반자동적인 방법을 통하여 의미 있는 패턴, 규칙, 관계 등을 찾아내는 것을 의미한다. 또는 데이터베이스로부터 지금까지 잘 알려지지 않는 유용하고 활용이 가능한 정보를 추출하는 과정으로 정의되기도 한다.(Noh, Kyoo-Sung, 2016)

기업이나 공공 기관에서는 일반적인 업무 활동을 통해 대용량의 데이터를 축적해왔으나 빈번하게 발생하는 많은 양의 데이터들은 수치화가 되지 않았을 뿐만 아니라, 수리적 형태보다는 질적인 형태로 저장이 되어 일반적인 통계 방법에 의해 분석과 활용이 될 수 없었다. 그러나 이러한 데이터베이스에는 발견하지 못한 패턴과 전략에 활용될 수 있는 정보들이 존재할 수 있기 때문에 데이터를 정제하고 가공할 필요성이 생겨났으며 이러한 데이터를 분석하여 기업에 필요한 자산으로 추출하는 과정이 데이터 마이닝이다.

데이터 마이닝은 정형 데이터 분석을 주로 다루는 데이터 마이닝과 비정형 데이터 마이닝으로 구분된다. 정형 데이터 마이닝을 일반적으로 데이터 마이닝이라 하며, 연관관계 분석, 군집 분석, 의사결정나무 기법, 인공신경망 기법, 사례기반추론 등이 이에 해당한다. 이와 대응되는 개념으로 정형화 되지 않은 데이터 또는 미리 정의된 데이터 모델을 가지고 있지 않은 데이터를 분석하는 것을 비정형 데이터 마이닝이라 지칭한다. 비정형 데이터의 유형에는

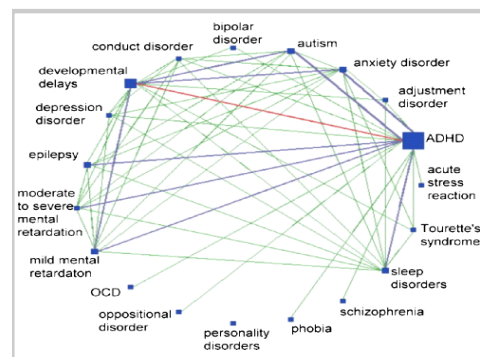
텍스트, 이미지, 음성과 영상, 로그 파일 등이 이에 해당되며, 비정형 데이터 마이닝 기법에는 텍스트 마이닝, 웹 마이닝, 오픈피언 마이닝, 소셜 네트워크 분석 등이 있다.

다양한 데이터 마이닝 방법이 존재하고 있으나 그 중에서 연관관계 분석, 군집 분석, 의사결정나무, 인공신경망 기법, 웹마이닝, 오픈피언 마이닝 등이 보편적으로 많이 사용되고 있으며, 본 연구에서는 이에 대한 분석과 연구모델에 적용을 위하여 이를 중심으로 연구를 진행하였다.

1. 데이터 마이닝

정형 데이터 마이닝의 연관관계 분석(Association Rule Mining)은 상품 혹은 서비스 간의 관계를 살펴보고 이로부터 유용한 규칙을 찾아내고자 할 때 이용될 수 있는 기법으로, 일반적으로 사건들은 동시 다발적으로 발생하며 이러한 사건들은 상호 영향을 주면서 결과를 나타나게 되는데 이와 같이 사건 또는 품목 간에 일어나는 연관성을 규명하려는 것이 연관성 규칙이다. 연관성 규칙은 데이터들의 빈도수와 동시 발생 확률을 이용하여 한 항목들의 그룹과 다른 항목들의 그룹 사이에 강한 연관성이 있음을 밝혀주는 기술이다. 예를 들어, (Item set X) → (Item set Y) 또는 (if X then Y : 만일 X가 일어나면 Y가 일어난다.)와 같이 표현할 수 있으며, Figure 1.은 연관관계 분석을 도식화한 예이다.

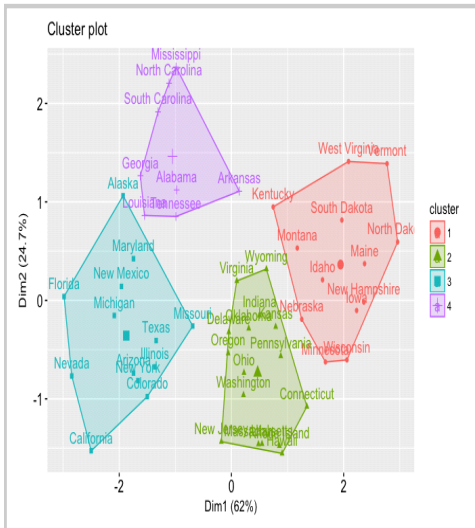
Fig. 1. Association Rule Mining graph sample



source : Tai, Yueh-Ming, and Hung-Wen Chiu. (2009)

군집 분석(Cluster Analytics)은 전체 데이터를 군집을 통해 잘 구분하는 것으로 다양한 특징을 가진 관찰 대상으로부터 동일 집단으로 분류하는데 사용하는 방식으로 개체들에 대한 사전 지식 없이 유사도에 근거하여 군집들을 구분하는 것이다. 개체 공간에 주어진 유한 개의 개체들이 서로 가깝게 모여서 무리를 이루고 있는 개체 집합을 군집(cluster)이라 부르며 군집화하는 과정을 클러스터링이라 칭한다. Figure 2.와 같이 군집 내의 개체들 간에는 유사도가 높으며, 이질적인 집단들 간에는 유사도가 동일 군집 내의 개체들보다 상대적으로 낮다는 특성을 갖는다. 군집 분석에는 퍼지군집, 혼합분포군집, 중심밀도군집, 격자기반군집 등 다양한 군집 분석이 존재하고 있으나 연구의 목적에 따라 일부만을 소개하였다.

Fig. 2. Cluster Analytic graph sample

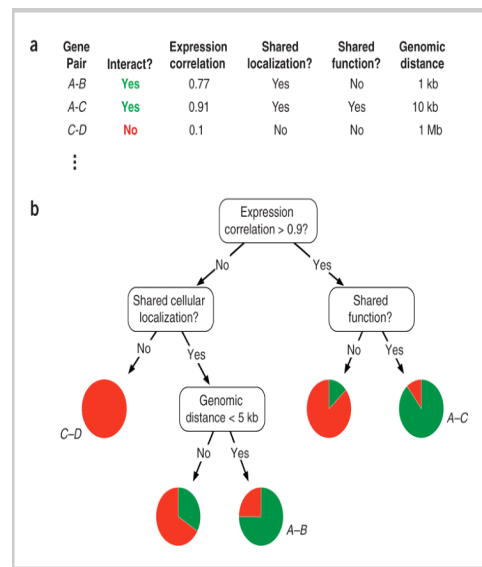


source : UC Business Analytics R Programming Guide

의사결정나무 기법(Decision Tree)은 데이터 마이닝의 주요 기법 중 하나로서 분류 및 예측에 주로 사용이 되는 기법이다. 이 기법은 경영, 경제에 관련된 다양한 분야의 예측에 이용이 되고 있으며, 사용이 비교적 용이하고 그 결과를 이해하기가 수월하다는 장점을 가지고 있다. 목표변수에 대한 의사결정 규칙(Rule)들을

나무 구조로 그래프화하여 분류와 예측을 수행하는 기법으로서 각종 데이터로부터 규칙을 도출하는데 매우 유용하게 활용되고 있으며 고객 분류, 기업의 부도 예측, 주가 예측, 환율 예측, 경제 전망 등 다양한 분야에서 활용이 되고 있다. Figure 3.과 같이 이진트리 구조로 도식화할 수 있으며 각각의 노드가 두 개의 자식노드를 만들어 yes 또는 no 질문에 답함으로써 터미널노드까지 진행해 나가는 방향으로 단순한 이진트리 모양만 있는 것이 아니라 혼합된 형태의 모형도 존재한다.

Fig. 3. Decision Tree graph sample

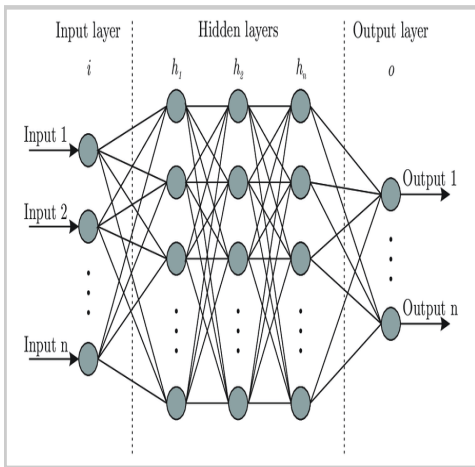


source : Kingsford, Carl and Steven L. Salzberg (2008)

인공신경망(Artificial Neural Network)기법은 생물학적 뇌의 작동 원리를 그대로 모방하는 방법으로 데이터 안의 독특한 패턴이나 구조를 인지하는데 필요한 모델을 구축하는 기법이다. 인공신경망은 간단한 계산 능력을 가진 처리 단위, 뉴런 또는 노드들이 서로 복잡하게 연결된 컴퓨터 시스템으로서 외부에서 주어진 입력에 대하여 반응할 수 있다. 이는 다수의 뉴런끼리의 상호연결성에 기인한 것으로 일반적인 인공신경망 모형으로는 Figure 4.와 같은 다계층 퍼셉트론 모형이 있다. 학습 패러다임에

근거한 인공신경망은 입력 데이터를 기초로 가중치를 통해서 의사결정을 하게 되는데, 인공신경망 모형은 예측 오차를 줄이고 예측 정확성을 증진시키기 위해서 반복적으로 가중치를 수정하며 이런 반복적인 단계를 훈련이라고 한다.

Fig. 4. Artificial Neural Network Architecture sample



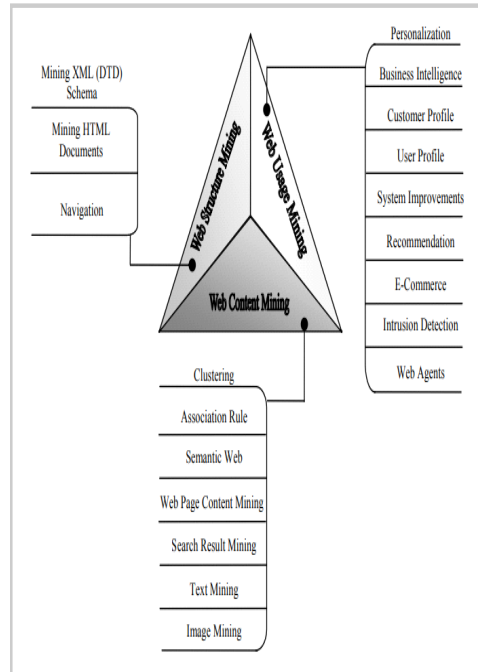
source : Bre, Facundo, Juan M. Gimenez and Victor D. Fachinotti (2018)

2. 비정형 데이터 마이닝

비정형 데이터 마이닝의 웹마이닝(Web Mining)은 인터넷을 통해 웹 서비스를 이용하면서 웹에서 패턴을 발견하는 것으로 웹 데이터의 속성이 반정형이거나 비정형이고, 링크(Link) 구조를 가지고 있기 때문에 비정형 데이터 마이닝으로 분류된다. 웹 마이닝은 Figure 5.와 같이 웹 사용 마이닝(Web Usage Mining), 웹 구조 마이닝(Web Structure Mining), 웹 콘텐츠 마이닝(Web Content Mining) 등으로 구분되며 웹 사용 마이닝은 웹 상에서 사용자가 찾고자 했던 것을 기록하고 있는 웹 서버 로그(Web Server Log)에서 유용한 정보를 추출하는 과정이다. 웹 구조 마이닝은 웹 구조 유형에 따라 하이퍼링크로부터 패턴을 추출하거나 HTML이나 XML과 같은 문서 구조를 분석한다. 웹 콘텐츠 마이닝은 웹페이지

지에서 유용한 데이터, 정보, 지식을 마이닝하고 출하여 통합하는 것을 의미한다.

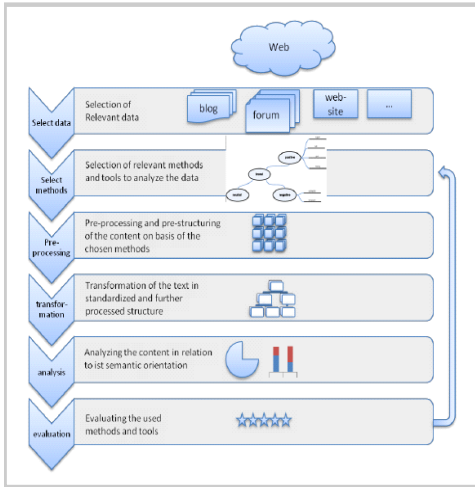
Fig. 5. Web Mining Architecture sample



source : Hassanzadeh, Hamed and Mohammad Reza Keyvanpour (2012)

오피니언 마이닝(Opinion Mining)은 어떤 사안이나 인물, 이슈, 이벤트 등과 같은 원천 데이터에서 의견이나 평가, 태도, 감정 등과 같은 주관적인 정보를 식별하고 추출하는 방법이다. 일반적으로 오피니언 분석은 문서, 문장, 특정, 관점 수준에서 표현된 견해가 긍정적인지, 부정적인지, 중립적인지 등을 주어진 텍스트의 특성을 분류하여 어떤 화제나 문서의 전반적 문맥 특성과 관련된 작성자나 화자의 태도를 파악하는데 도움을 준다. 주요 분석 대상은 포털 게시판, 블로그, 쇼핑몰과 같은 대규모의 웹 문서이기 때문에 자동화된 분석 방법을 사용하고, 감정분석이라 불리며 트렌드 파악, 제품·서비스 평가, 미래 예측 등에 활용되어지고 있다. Figure 6.는 오피니언 마이닝 기법의 과정을 도식화한 것이다.

Fig. 6. Process Model for Opinion Mining sample



source : Petz, Gerald, et al. (2012)

III. 연구모형의 설계

1. AHP 연구방법

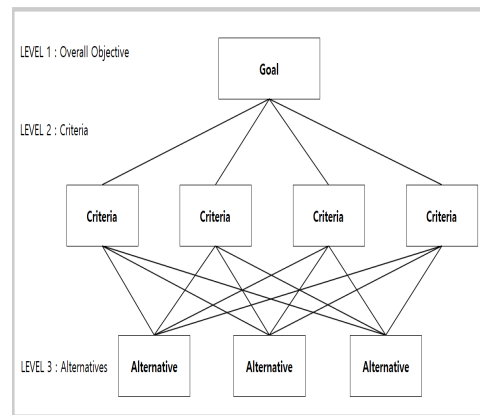
본 연구는 AHP 분석방법을 활용하였으며, 계층화 분석과정(Analytic Hierarchy Process, AHP)은 복잡한 의사결정 문제를 계층적으로 표현하고 그 계층의 항목들 간의 쌍대비교(Pairwise comparison)를 통하여 최선의 대안을 도출해 내는 의사결정 기법이다(Saaty, 1988). 일반적으로 비교 척도가 다르거나, 척도가 없는 경우의 정량적 분석기법의 한계를 보완하고, 다수의 의사결정의 목표 또는 평가기준이 복합적인 경우에 대안들의 체계적 평가를 위하여 활용되고 있다.

AHP 기법을 통한 다기준 의사결정 분석절차는 일반적으로 첫째 계층구조 형성, 둘째 쌍대비교의 수행, 셋째 가중치 산출, 넷째 일관성 평가, 그리고 마지막으로 최종대안 선정의 다섯 단계를 거치면서 수행된다.

계층구조 형성은 상단의 Level 1에 목표(Goal)를 설정하고 Level 2에 평가기준(Criteria)이 되는 요인들을 구성한다. 하단의

Level 3의 경우 두 가지 방식이 가능하며 평가기준을 기반으로 대안들(Alternatives)을 구성하여 의사결정의 우선순위를 구할 수도 있고, Level 2의 평가기준에 대하여 하단의 세부적인 요인들(Factors)로 구성하여 요인들의 중요도(Importance)를 산출할 수도 있다. 본 연구에서는 Figure 7.와 같이 Level 3를 대안들로 구성하였다.

Fig. 7. Alternative decision making model



둘째, 평가기준들 간에 쌍대비교를 수행하고 평가기준 하에 대안들 간의 쌍대비교를 수행한다. 쌍대비교에 사용하는 척도는 일반적으로 양방향 1에서 9까지의 비율척도로서 각 수치의 의미는 Table 1.과 같다. 쌍대비교의 결과 값은 AHP 분석을 위한 기초자료로 활용된다.

Table 1. Pairwise comparison Scale for AHP

Scale	Degree of preference
1	Equal Importance
3	Moderate importance of one factor over another
5	Strong or essential importance
7	Very strong or demonstrated importance
9	Extreme importance
2,4,6,8	Values for inverse comparison

Table 2. Analytic Hierarchy Criteria and Alternatives

factor	description
Association Rule Mining	discovering interesting relations between variables in large databases
Cluster Analytics	grouping a set of objects in such a way that objects in the same group
Decision Tree	decision support tool and their possible consequences, including chance event outcomes, resource costs, and utility
Artificial Neural Network	computing systems inspired by the biological neural networks that constitute animal brains.
Web Mining	data mining techniques to discover patterns from the World Wide Web
Opinion Mining	use of natural language processing, text analysis, computational linguistics, and biometrics
Transport Cooperation	Sharing equipment or vehicles for transportation between businesses
Distribution Center Cooperation	A form of joint use of storage facilities such as warehouses and distribution centers
Logistics Information System Cooperation	Joint use of a system that processes cargo information in logistics activities
Logistics Partnership Cooperation	Logistics companies form cooperative relationships to share logistics services and infrastructure

셋째, 평가기준과 대안들에 대한 가중치 (Weighted values)를 산출하기 위해서는 쌍대 비교행렬(Pairwise comparison Matrix)을 정리 하여 정규화(Normalization)하고 정규화 행렬의 각 행의 평균을 구한다. 각 행의 평균이 가중치이며, 가중치의 합은 1이 된다.

넷째, 일관성(Consistency) 평가는 의사결정자가 수행한 쌍대비교의 논리적인 모순을 측정하는 것으로 대안 선정에 앞서 수행된다. 일관성 평가는 일관성 비율(Consistency Ratio, CR)을 측정함으로써 쌍대비교의 일관성 여부를 판단한다. 일관성 비율은 일반적으로 0.1이하의 기준을 적용할 경우 합리적인 수치, 0.2이하일 경우에는 허용할 수 있는 수치로 평가된다. 일관성 비율을 계산하는 순서는 다음과 같다.

1단계 : 일관성 수치인 λ_{max} 를 계산한다. λ_{max} 의 계산은 쌍대비교행렬의 행들과 평균값들을 곱한 행렬에 평균값들로 행렬 나눗셈을 한

후 최종적으로 행렬 값들의 평균을 산출한다.

2단계 : 일관성 지수 (Consistency Index, CI)를 계산한다. CI를 산출하기 위한 식은 아래와 같으며, 여기서 n 은 비교되는 항목의 수를 의미한다.

$$CI = \frac{\lambda_{max} - n}{n - 1}$$

3단계 : 일관성 비율 (Consistency Ratio, CR)을 계산한다. CR를 산출하기 위한 식은 아래와 같이 이전 단계에서 산출한 CI를 임의지수(Random Index, RI)로 나누어 생성한다. RI는 무작위 샘플링 실험을 통하여 제시한 수치들로 RI 매트릭스에서 요인의 수에 해당하는 값을 취한다.

$$CR = \frac{CI}{RI}$$

마지막으로 의사결정자가 수행한 쌍대비교가 일관성에 문제가 없는 경우 최종대안은 가중치 산출을 통해 얻은 평가기준들의 상대적

Fig. 8. Big Data Mining AHP Model

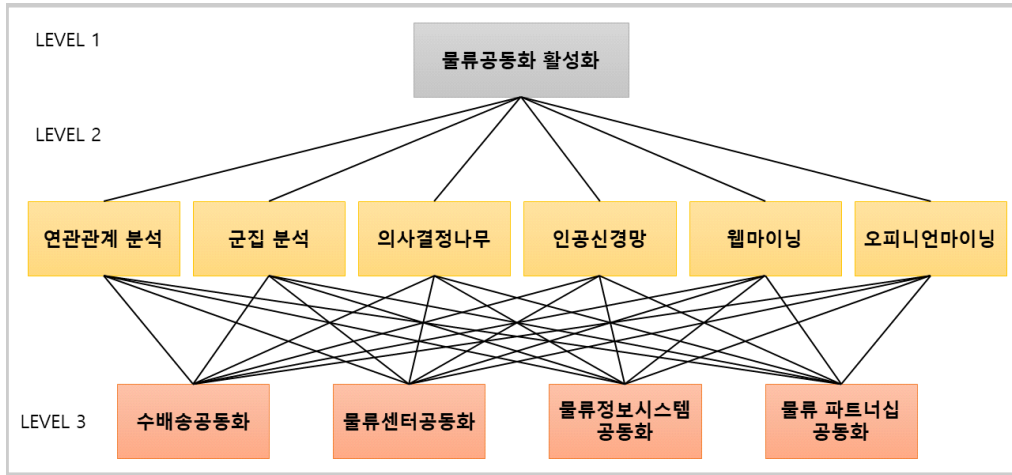


Table 2. Analytic Hierarchy Criteria and Alternatives

factor	description
Association Rule Mining	discovering interesting relations between variables in large databases
Cluster Analytics	grouping a set of objects in such a way that objects in the same group
Decision Tree	decision support tool and their possible consequences, including chance event outcomes, resource costs, and utility
Artificial Neural Network	computing systems inspired by the biological neural networks that constitute animal brains.
Web Mining	data mining techniques to discover patterns from the World Wide Web
Opinion Mining	use of natural language processing, text analysis, computational linguistics, and biometrics
Transport Cooperation	Sharing equipment or vehicles for transportation between businesses
Distribution Center Cooperation	A form of joint use of storage facilities such as warehouses and distribution centers
Logistics Information System Cooperation	Joint use of a system that processes cargo information in logistics activities
Logistics Partnership Cooperation	Logistics companies form cooperative relationships to share logistics services and infrastructure

중요도와 각 평가기준 하에서 대안들의 선호도를 곱한 값들 중 가장 높은 수치를 나타내는 대안이 선택된다.

2. AHP 조사목적 및 방법

본 연구에서는 물류공동화 결정으로 인한 활성화를 목적으로 하며, 물류공동화를 결정하기 위한 데이터 수집 및 의사결정을 위한 정보를

Table 3. R functions for AHP analysis

function name	description
f_num()	function that makes the data value into fractions
f_pw_matix()	function that builds pairwise comparison matrix from data
f_n_matix	function to perform normalization of a matrix
f_ave	function for calculating weighed values
f_priority	function to perform normalization of a matrix and to calculate weighed values
f_consistency	function for calculating a consistency ratio
f_geometric.mean	function that makes data geometric
f_consistency_list	function that produces a consistency ratio list
f_compatibility_list	function that produces a compatibility index list
f_ahp_main	function to build pairwise comparison matrix, to perform normalization of the matrix and to calculate weighed values for all data

Table 4. CR(Consistency Ratio; CR)

idv	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_10	P_11	P_12
1	0.179	0.083	0.097	0.174	0.083	0.086	0.087	0.142	0.076	0.189	0.089	0.109
2	0.188	0.047	0.142	0.000	0.047	0.190	0.082	0.089	0.108	0.128	0.047	0.133
3	0.117	0.097	0.096	0.105	0.146	0.092	0.095	0.102	0.119	0.115	0.101	0.097
4	0.117	0.001	0.121	0.026	0.000	0.170	0.087	0.066	0.067	0.126	0.012	0.127
5	0.162	0.074	0.185	0.064	0.000	0.120	0.053	0.083	0.073	0.075	0.064	0.174
6	0.051	0.080	0.127	0.147	0.102	0.030	0.051	0.107	0.051	0.091	0.070	0.121
7	0.053	0.076	0.133	0.120	0.159	0.073	0.053	0.092	0.131	0.066	0.096	0.084

제공하는 효율적인 데이터 마이닝(연관관계 분석, 군집 분석, 의사결정나무, 인공신경망, 웹마이닝, 오피니언 마이닝)의 우선순위를 분석하기 위하여 평가기준을 데이터 마이닝으로 설정하였다. 대안들은 데이터 마이닝을 기반으로 선택 가능한 물류공동화 유형(수배송공동화, 물류센터공동화, 물류정보시스템 공동화, 물류파트너십 공동화)을 세분화하여 중요도를 산출할 수 있도록 구성하였다. Table 2. 는 본 연구에서 활용된 요인별 정의를 정리하였고, Figure 8.은 이를 가시적으로 모형화하였다.

평가 기준은 데이터 마이닝 기법으로 조사 대상은 빅데이터 분야 및 물류 분야 전문가 12명을 대상으로 2021년 9월 한 달 간 설문지를 수집하였으며, 설문 대상은 모두 학계의 교수 및 박사급 연구자와 연구소의 연구원으로 구성

하였다. 물류 분야 교수 4명, MIS 박사급 연구자 3명, 빅데이터 전문가 5명을 대상으로 설문조사 한 결과치를 AHP 분석방법으로 반영하였다. 분석도구로는 빅데이터 분석이 용이하고 오픈소스로 운영되고 있는 R(x64 4.1.1)을 활용하였으며, 작업 환경은 R Studio를 이용하였다. 본 연구에서는 AHP 분석을 위한 기능들에 대하여 Table 3.와 같은 사용자 R 함수(Yoon, Cheol-Ho, Choi, Kwang-Don, 2019)들을 사용하였다. R 함수들의 세부내용은 APPENDIX A에 첨부하였다.

아래의 사용자 R 함수들은 AHP 분석을 위하여 R파일(ahp_functions.R)로 생성하여 작업영역 Directory에 보관하였다. R 패키지로는 ‘기하평균함수’ 및 ‘list 작성기능’을 설치하였다.

```
#AHP 분석시 활용되는 R 패키지 설치
install.packages("psych") #기하평균함수
install.packages("rlist") #list작성기능
```

분석을 위하여 자료는 CSV파일(ahp_mining_data.csv)로 생성하여 작업영역에 보관하였으며 R로 자료를 불러들이기 위해 아래의 명령어를 실행하였다.

```
source("ahp_functions.R") #AHP함수
ahp_data<-read.csv("ahp_mining_data.csv"
, header = T) #자료 R로 로드
ahp_data #자료보기
```

본 연구의 모델은 평가기준이 6개이고 그 하위 대안이 각 각 4개씩으로 구성된 계층구조를 가지고 있으며 쌍대비교 수는 평가기준에 대한 쌍대비교문항 15개와 대안에 대한 쌍대비교문항 36개로 총 51개의 응답문항들로 구성되었다.

```
#평가기준 및 대안수 설정
hs = c(6,4,4,4,4,4,4)
```

AHP 분석은 자료의 유효성을 확인하기 위하여 응답자별 일관성 평가가 이루어져야 하며 함수 f_consistency_list()를 이용하여 수행한다.

```
#응답자별 일관성 비율 산출
consistency_list<-f_consistency_list(ahp_data, hs)
consistency_list
```

일관성 비율은 0.1 이하가 권장되지만 0.2 이하일 경우에는 허용할 수 있는 수치로 평가하였다. 본 연구에서 응답자 12명의 일관성 비율(CR)은 Table 4.와 같다. CR이 0.2 이상되는 항목을 포함하는 경우 해당 자료를 분석에서 제외해야 되며 이를 위한 명령어는 아래와 같다.

```
# 일관성 검정 탈락 항목들
consistency_fails<-subset(consistency_list,
CR >0.2)
# 일관성 미비 제외응답자
fails_id=as.vector(unique(consistency_fails
$p_man))
cr_data=ahp_data
for(i in 1:length(fails_id)) {
  cr_data=subset(cr_data,id !=fails_id[i])
}
```

```
cr_data
```

IV. 연구 결과 및 분석

1. 통합된 항목별 가중치 산출

통합 항목별 중요도 산출은 일관성과 대응가능지수 검정을 통과한 응답자들의 의견을 통합하여 AHP 분석을 위한 통합된 항목별 가중치를 산출하는 절차로 최종 대안을 산출하기 위한 전단계에 해당된다.

일관성 및 대응가능지수 평가를 통과한 자료인 si_data를 가지고 통합자료 가중치 산출 함수인 f_ahp_main()을 이용하여 산출한다.

```
# 통합자료 가중치 산출
ave_main<-f_ahp_main(si_data, hs)
ave_main$ave_list
```

Fig. 9. Weighed Values

```
> ave_main$ave_list
  class options      ave
1      1      1 0.36128571
2      1      2 0.12030507
3      1      3 0.08660247
4      1      4 0.10380312
5      1      5 0.23778490
6      1      6 0.09021872
7      2      1 0.18826881
8      2      2 0.15143797
9      2      3 0.44806599
10     2      4 0.21222723
11     3      1 0.14452642
12     3      2 0.40538106
13     3      3 0.27975012
14     3      4 0.17034239
15     4      1 0.09544548
16     4      2 0.10327132
17     4      3 0.47199540
18     4      4 0.32928780
19     5      1 0.11291613
20     5      2 0.14213115
21     5      3 0.49539308
22     5      4 0.24955964
23     6      1 0.15387962
24     6      2 0.21682634
25     6      3 0.45816488
26     6      4 0.17112916
27     7      1 0.42067546
28     7      2 0.16250210
29     7      3 0.28959291
30     7      4 0.12722953
```

Fig. 10. Criteria Weighed Values

```

> ave_main$ahp_list
$class
[1] 1

$m
      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
[1,] 1.0000000 3.9017080 4.545463 3.0613823 1.1477228 4.9660971
[2,] 0.2562980 1.0000000 1.462865 1.0455885 0.4054801 2.3458926
[3,] 0.2199996 0.6835899 1.000000 0.8491907 0.4277147 1.0626586
[4,] 0.3266498 0.9563992 1.177592 1.0000000 0.5867760 0.6135262
[5,] 0.8712906 2.4662121 2.338007 1.7042279 1.0000000 2.8172691
[6,] 0.2013654 0.4262770 0.941036 1.6299222 0.3549537 1.0000000

$nm
      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
[1,] 0.34775310 0.41357123 0.39646554 0.32952417 0.2925888 0.38781141
[2,] 0.08912843 0.10599748 0.12759442 0.11254612 0.1033690 0.18319495
[3,] 0.07650555 0.07245881 0.08722226 0.09140605 0.1090373 0.08298491
[4,] 0.11359349 0.10137591 0.10271222 0.10763902 0.1495867 0.04791136
[5,] 0.30299400 0.26141228 0.20392626 0.18344142 0.2549299 0.22000558
[6,] 0.07002543 0.04518429 0.08207929 0.17544322 0.0904883 0.07809179

$ave
[1] 0.36128571 0.12030507 0.08660247 0.10380312 0.23778490 0.09021872

$nm_ave
      ave rank
[1,] 0.34775310 0.41357123 0.39646554 0.32952417 0.2925888 0.38781141 0.36128571 1
[2,] 0.08912843 0.10599748 0.12759442 0.11254612 0.1033690 0.18319495 0.12030507 3
[3,] 0.07650555 0.07245881 0.08722226 0.09140605 0.1090373 0.08298491 0.08660247 6
[4,] 0.11359349 0.10137591 0.10271222 0.10763902 0.1495867 0.04791136 0.10380312 4
[5,] 0.30299400 0.26141228 0.20392626 0.18344142 0.2549299 0.22000558 0.23778490 2
[6,] 0.07002543 0.04518429 0.08207929 0.17544322 0.0904883 0.07809179 0.09021872 5
    
```

통합자료 가중치 산출인 Figure 9.의 값들은 평가기준 및 대안수 설정시 사용된 순서 6,4,4,4,4,4 항목의 순서이다. 즉 처음 6개는 모형의 평가기준인 연관관계 분석, 군집 분석, 의사결정나무, 인공지능망, 웹마이닝, 오피니언 마이닝의 상대적 가중치(또는 중요도) 값들이고 나머지 4개씩은 각 평가기준 영역에서 획득한 물류공동화 세부유형들(수배송공동화, 물류센터공동화, 물류정보시스템 공동화, 물류 파트너십 공동화)의 상대적 가중치 값들이다.

통합된 항목별 가중치 산출에 활용된 항목별 쌍대비교 행렬, 정규화 행렬, 가중치 값들은 아래와 같은 명령어 수행으로 확인할 수 있다.

```

# 전체[쌍대비교 행렬][정규화 행렬][가중치] 산출
ave_main$ahp_list
    
```

Figure 10.은 본 연구의 평가기준의 가중치 산출 결과로 평가기준 1번 연관관계 분석이 우선순위가 가장 높은 것을 확인할 수 있다. 2 순위로는 비정형 데이터 마이닝 기법인 웹 마이닝의 우선 순위가 높았으며 3순위는 군집 분석으로 평가되었다. 다음으로는 인공지능망, 오피니언 마이닝, 의사결정나무 순으로 평가된 것을 확인할 수 있다. (1순위: 연관관계 분석, 2순위: 웹 마이닝, 3순위: 군집 분석, 4순위: 인공지능망, 5순위: 오피니언 마이닝, 6순위: 의사결정나무)

2. 최종대안의 선택

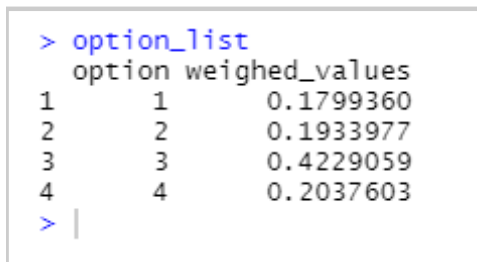
최종 대안의 선택은 제안서별 산출된 가중치와 평가기준별로 가중치의 결합을 통하여 이루어진다. 따라서 다음과 같은 명령어로 제안서

들에 대한 최종 가중치를 산출한다.

```
aves=ave_main$ave_list$ave
option_list = NULL
for(i in 1:4) {
  wv = aves[1] * aves[i+6] + aves[2] *
aves[i+10] + aves[3] * aves[i+14] + aves[4] *
aves[i+18] + aves[5] * aves[i+22] + aves[6] *
aves[i+26]
  option_list=rbind(option_list,
data.frame(option=i, weighed_values=wv))
}
option_list
```

Figure 11.은 본 연구의 최종 가중치 산출 결과로 대안 3번-물류정보시스템공동화가 가장 높은 수치를 보여주고 있다. 따라서 데이터 마이닝을 활용하여 물류공동화를 실천하는 경우 물류정보시스템 공동화가 가장 우선 시 되며, 다음으로 물류 파트너십 공동화, 물류센터공동화, 수배송공동화 순으로 우선 순위가 선정된 것을 알 수 있다.

Fig. 11. Final Weights



option	weighed_values
1	0.1799360
2	0.1933977
3	0.4229059
4	0.2037603

본 연구결과 데이터 마이닝은 연관관계 분석 및 웹 마이닝, 군집분석 순으로 중요성이 나타났으며 이를 기반으로 물류공동화 활성화에 있어서 연관관계 분석의 활용 가능성이 가장 높은 것으로 판단할 수 있다. 이외에 웹 마이닝과 군집분석이 중요한 것으로 결과가 나타났으며, 이후 빅데이터 전문가 면담을 통해 확인한 바에 따르면 웹 콘텐츠 마이닝을 통한 기존 물류 데이터 뿐만이 아닌 물류공동화에 유용한 타 데이터 및 정보를 추출하는 것으로 물류공동화

활성화에 대한 긍정적인 요소를 평가하였다. 대안적인 측면에 있어서는 물류정보시스템이 가장 높은 수치를 보인데 있어, 데이터 마이닝의 특성 상 정보시스템 측면이 강조되었으며 이에 대한 활용 가능성이 가장 높은 것으로 여겨지며, 물류 파트너십의 우선 순위가 차 순위인 이유에 대해서는 파트너 정보에 대한 데이터 마이닝 활용이 다양하기 때문인 것으로 분석하였다. 그 다음으로 물류센터공동화와 수배송공동화에 있어서는 당위적으로 중요하나 데이터 마이닝이라는 측면을 반영하여 우선순위가 낮게 나왔을 뿐 가중치 측면에 있어서는 물류파트너십 공동화, 물류센터공동화, 수배송공동화의 차이는 크지 않은 것으로 분석하였다.

V. 결론

4차 산업혁명 시대의 기술 발전은 다양한 산업의 패러다임을 변경하고 있다. 빅데이터를 비롯하여 클라우드, 인공지능, 가상현실, 사물인터넷 등 다양한 기술들이 활용되며 기존의 산업과 시너지 효과들을 내며 급진적인 발전과 가치 창출을 해내고 있다. 그 중에서 물류 분야는 기존부터 정량적인 데이터에 영향을 많이 받았으며 지속적으로 데이터를 축적하고 관리를 해왔기에 빅데이터 분석과 연계 가능성이 높고 활용 효과도 높다. 빅데이터가 가지는 특성에 따라 단순히 대규모 데이터를 기존의 방식으로 활용하는 것만으로도 충분히 유의미한 결과를 가져오는 경우도 있다. 그러나 현대의 발달된 기술은 이러한 빅데이터에서 숨겨진 패턴과 새로운 상관관계를 발견하는 데이터 마이닝 기술이 함께 발달하였으며 이를 통해 전에 없던 유의미한 결과물들을 도출해내고 있다. 그러므로 빅데이터 분석에서 데이터 마이닝은 중요한 부분을 차지하고 있으며 본 연구에서는 이러한 데이터 마이닝 기술을 활용하여 물류분야 및 물류공동화에 기여할 수 있는 데이터 마이닝 기법을 분석하고자 하였다.

이에 AHP 기법을 활용하여 물류공동화에 효율적인 데이터 마이닝을 유형별로 우선순위를 도출하고자 하였으며, 이를 분석하는 도구는

기존의 Expert Choice가 아닌 R 프로그램과 R Studio를 활용하였다. R은 오픈소스로 누구나 활용이 가능하며, 본 연구에서는 사용한 함수를 모두 밝히고 있어 이를 활용하고자 하는 후속 세대의 연구에도 기여하고자 하였다. 본 연구에서 활용된 AHP 함수는 다른 AHP 연구에도 활용될 수 있으며, 이를 통하여 다양한 후속 연구가 이루어질 수 있기를 바란다.

연구의 결과는 평가기준(Criteria)의 우선순위에 있어서는 연관관계 분석, 웹 마이닝, 군집 분석, 인공신경망, 오피니언 마이닝, 의사결정 나무 순으로 평가되었다. 결과를 기반으로 물류 분야 전문가와 인터뷰를 진행하였으며, 물류 분야의 특성 상 인과관계를 증시하고 직접적인 상관관계에 비중을 많이 두기 때문에 연관관계 분석이 가장 우선순위가 높은 것으로 판단하였다. 2순위로 평가된 웹 마이닝의 경우, 기존의 물류 분야에 활용되는 데이터가 대다수 정형 데이터임에도 불구하고 비정형 데이터 마이닝인 웹 마이닝이 높은 순위를 차지하여 비정형 데이터의 활용에 대한 연구가 더 활발히 이루어져야 함을 시사하였다.

다음으로 대안들(Alternatives)의 우선순위에 있어서는 물류정보시스템 공동화, 물류파트너십 공동화, 물류센터공동화, 수배송공동화 순으로 평가되었다. 이는 데이터 마이닝의 종합적인 평가에 따라 정해진 우선순위이며, 각 각의 데이터 마이닝 기법에 따라 적용해야 하는 물류공동화 세부 유형은 각 각 다른 결과가 도출되었다. 즉, 전반적인 대안으로 선택의 참고 사항은 될 수 있으나 물류공동화 유형을 선택하는 절대적인 대안이 될 수 없음을 의미한다.

본 연구결과를 기반으로 물류공동화 활성화

를 위하여는 기본적으로 물류의 기본 핵심을 파악하는데 있어서 연관관계 분석이 그 중요도가 가장 높은 것으로 분석되었고, 웹 마이닝이 높은 우선 순위를 보인 이유는 웹 콘텐츠 마이닝을 통한 기존 물류데이터 뿐만이 아닌 물류공동화에 유용한 타 데이터 및 정보를 추출하는 것이 물류공동화에 있어서 긍정적인 요소로 평가되었다. 또한 대안들의 우선 순위에 있어 물류정보시스템 공동화에 있어서 데이터 마이닝 활용 가능성이 가장 높으며 이로 인한 시너지 효과가 높을 것으로 분석되었다. 이외의 물류파트너십 공동화, 물류센터공동화, 수배송공동화는 가중치에 있어서 우선순위는 결정되었으나 그 값 자체에는 큰 차이가 없는 것으로 분석하였다. 따라서 물류공동화의 활성화를 위해 연관관계 분석 및 웹 마이닝, 군집 분석 등을 활용할 필요성이 강조되며, 물류정보시스템의 공동화가 데이터 마이닝을 통한 물류공동화 활성화에 가장 큰 긍정적인 요인으로 작용할 것으로 여겨진다.

본 연구는 빅데이터 분야에 있어서 데이터 마이닝의 중요성을 설명하고, 분석 도구 측면에서 기존의 Expert Choice에서 벗어나 오픈소스로 공급되는 R을 활용하였다는 측면에서 유의미한 연구방법을 제시했다고 여겨진다. 그러나 이는 분석 도구 측면에서 차이이고, 분석 방법적인 측면에서는 기존의 AHP 기법을 활용하였다는 점에서 한계가 존재한다. 따라서 후속 연구에 있어서는 데이터 마이닝 또는 비정형 데이터 마이닝 자체를 활용한 다양한 연구가 이루어져 무역과 물류 분야에 있어서 빅데이터 활용 범위가 점차 확대되어지기를 바란다.

References

- Bre, Facundo, Juan M. Gimenez and Victor D. Fachinotti (2018) "Prediction of wind pressure coefficients on building surfaces using artificial neural networks.", *Energy and Buildings*, 158, pp. 1429-1441.
- Hassanzadeh, Hamed and Mohammad Reza Keyvanpour (2012), "Semantic web requirements through web mining techniques." *International Journal of Computer Theory and Engineering*, 4(4), pp. 616-620.

- Kim, Seung-Wook (2017), "Value Model for Applications of Big Data Analytics in Logistics" *Journal of Digital Convergence*, 15(9), pp. 167-178.
- Kingsford, Carl and Steven L. Salzberg (2008), "What are decision trees?" *Nature biotechnology*, 26(9), pp. 1011-1013.
- Lee, Eun-Jeong (2021), "A Study on Application and Utilization method of Smart Logistics", *The Korean Academy for Trade Credit Insurance*, 22(1), pp. 69-84. DOI : 10.22875/jiti.2021.22.1.005
- Lee, Sang-Won, Kim Sung-Hyun (2016), "Finding Industries for Big Data Usage on the Basis of AHP", *Journal of Digital Convergence*, 14(7), pp. 21-27.
- Noh, Kyoo-Sung et. al. (2016), *Big data : Analyse data*, KASOM, Dept. of Bigdata, Gyeonggi-do, Korea.
- Park Kwang-O, Jung Dae-Hyun, Kwon Sang-Min. (2020), "Necessity of the Physical Distribution Cooperation to Enhance Competitive Capabilities of Healthcare SCM -Bigdata Business Model's Viewpoint-", *Management & Information Systems Review*, 39(3), pp. 17-35.
- Petz, Gerald, et al. (2012), "On text preprocessing for opinion mining outside of laboratory environments." *International Conference on Active Media Technology*. Springer, Berlin, Heidelberg, pp. 618-629.
- Saaty, Thomas L. (1988), "What is the analytic hierarchy process?", *Mathematical models for decision support*. Springer, Berlin, Heidelberg, pp. 109-121.
- Shin, Na-la and Kim, Yong-Jin (2018), "Activating Strategies of Logistics Big Data through Case Studies" *Korea Logistics Review*, 28(2), pp. 91-101.
- Tai, Yueh-Ming and Hung-Wen Chiu (2009), "Comorbidity study of ADHD: applying association rule mining (ARM) to National Health Insurance Database of Taiwan." *International journal of medical informatics*, 78(12), pp. 75-83.
- UC Business Analytics R Programming Guide, *K-means Cluster Analysis*. Available from https://uc-r.github.io/kmeans_clustering (accessed September 30, 2021)
- Yoon, Cheol-Ho, Choi, Kwang-Don (2019), "A Study on the Analytic Hierarchy Process of Group Decision Making Using R" *Journal of Information Technology and Architecture*, 16(4), pp. 405-418.

APPENDIX A

```

#-----
# 데이터 값 분수화 함수
#-----
f_num<- function(v) {
  f_v = c()
  for(i in 1:length(v)){
    if (v[i] > 0)
      f_v = append(f_v, v[i])
    else
      f_v = append(f_v, 1/v[i] * -1)
  }
  f_v
}
#-----
# 데이터 값 쌍대비교행렬화
#-----
f_pw_matrix<- function(v, cnt) {
  v = f_num(v) # 데이터 값 분수화
  r_v = 1/v # 상대 값 추가
  m = diag(1, cnt, cnt)
  idx =1
  for(i in 1:cnt) {
    j = i +1
    while(j <= cnt) {
      m[i, j] = v[idx]
      idx = idx +1
      j = j+1 }
    idx =1
    for(i in 1:cnt) {
      j = i +1
      while(j <= cnt) {
        m[j, i] = r_v[idx]
        idx = idx +1
        j = j+1 }
      }
    }
  m
}
#-----

# 정규화 행렬 함수
#-----
f_n_matrix<- function(m) {
  nm = t(t(m)/colSums(m))
  nm
}
#-----
# 중요도 산출 함수
#-----
f_ave <- function(nm) {
  ave = apply(nm, 1, mean)
  ave
}
#-----
# 정규화행렬 및 중요도 산출 함수
#-----
f_priority <- function(nm) {
  ave = apply(nm, 1, mean)
  rank = rank(-ave)
  nm_ave = cbind(nm,ave, rank)
  nm_ave
}
#-----
# 일관성 비율 산출 함수
#-----
f_consistency <- function(m, cnt) {
  nm <- f_n_matrix(m)
  ave <- f_ave(nm)
  # RI
  vRI = c(0.00, 0.00, 0.58, 0.9, 1.12, 1.24, 1.32,
  1.41, 1.46, 1.49)
  # 람다 max 산출
  lambda_max = mean((m %*% ave)/ave)
  #CI, CR 산출
  CI = (lambda_max - cnt)/(cnt-1)
  CR =CI/vRI[cnt]
  consistency =
  data.frame(lambda_max=lambda_max,
  CI=CI, CR=CR)
  consistency
}
#-----
# 자료의 기하평균화

```



```

#-----
f_geometric.mean <- function(data) {
library(psych)
data = apply(data, 2, f_num) # 분수화
if (nrow(data) < 2)
vars = data
if (nrow(data) > 1)
vars = geometric.mean(data) else #기하평균
함수
vars = data
vars
}
#-----
# 응답자별 일관성 지수 산출
#-----
f_consistency_list <- function(data, hs) {
p_man = data$id # 평가자 전체 ROW_ID
consistency_list = NULL
for(p_id in 1:length(p_man)) {
p_man_id = p_man[p_id]
vars = as.numeric(subset(data, id==p_man_id
)) #데이터 값
vars = vars[-1]
idv = 1
to = 0
for(idv in 1:length(hs)) {
cnt = hs[idv] # 항목수
v_count = hs[idv]*(hs[idv]-1)/2 # 측정지표수
from = to + 1
to = v_count + from - 1
v = vars[c(from:to)]
m <- f_pw_matrix(v, cnt) # 쌍대비교행렬
consistency <- f_consistency(m, cnt) # 일관
성 비율
산출
consistency_list <- rbind(consistency_list,
data.frame(p_man=p_man[p_id], idv,
CR=round(consistency$CR,3)))
}
}
consistency_list
}
#-----

```

```

# 응답자별 대응가능 지수 산출
#-----
f_compatibility_list <- function(data, hs) {
geo_data <- f_geometric.mean(data[-1]) # 기하
평균데이터
p_man = data$id # 평가자 전체
ROW_ID
s_index_list = NULL
for(p_id in 1:length(p_man)) {
p_man_id = p_man[p_id]
vars = as.numeric(subset(data, id==p_man_id
))
# 데이터 값
vars = vars[-1]
#-----
idv = 1
to = 0
for(idv in 1:length(hs)) {
cnt = hs[idv] # 항목수
v_count = hs[idv]*(hs[idv]-1)/2 # 측정지표수
from = to + 1
to = v_count + from - 1
v = vars[c(from:to)]
geo = geo_data[c(from:to)]
sm <- f_pw_matrix(geo, cnt) #기준행렬
m <- f_pw_matrix(v, cnt) #쌍대비교행렬
tm = t(m) #역비교행렬
cm = sm * tm
# 대응가능지수행렬
c_sum = apply(cm, 1, sum)
# 합계 산출 (1:row, 2:col)
s_index = sum(c_sum)/(nrow(m)*ncol(m))
# 대응가능지수
s_index_list = rbind(s_index_list,
data.frame(p_man=p_man[p_id], idv=idv,
s_index=s_index))
}
}
s_index_list
}
#-----
# 전체 [쌍대비교 행렬][정규화 행렬][중요도 지
수] 산출

```

```

#-----
library(rlist)
f_ahp_main <- function(data, hs) {
vars <- f_geometric.mean(data[-1]) # 기하평균
데이터
ave_list = NULL
ahp_list = list()
#-----
idv =1
to = 0
for(idv in 1:length(hs)) {
cnt = hs[idv] # 항목수
v_count = hs[idv]*(hs[idv]-1)/2 # 측정지표수
from = to +1
to = v_count + from - 1
v = vars[c(from:to)]
m <- f_pw_matrix(v, cnt) # 쌍대비교행렬
nm <- f_n_matrix(m) # 정규화행렬
ave <- f_ave(nm) # 평균값산출
ave_list = rbind(ave_list,
data.frame(class=idv,
options=c(1:cnt), ave=ave))
nm_ave <- f_priority(nm) # 정평균값 추가
ahp_list <-list.append(ahp_list, class=idv,
m=m,
nm=nm, ave=ave, nm_ave=nm_ave)
}
ave_list <- list(ave_list=ave_list,
ahp_list=ahp_list)
}

```