

건축공사감리 문서 기반 연관규칙 및 비용효율성 분석 모델

A Study on Association Rule and Cost Efficiency Analysis Model Using Construction Supervision Reports

송태근¹ · 유위성^{2*}

Song, Tae-Geun¹ · Yoo, Wi Sung^{2*}

Abstract : To improve the cost performance of construction sites, various systems and standards are constantly being developed and implemented. Although legal requirements for these system and standard improvements have been increasing, the cost efficiency performance of construction sites remains stagnant. We have digitized documents generated through construction supervision work at 39 building construction sites and proposed a model that can support decision-making in cost efficiency evaluation. This model selects key keywords that are considered to be highly related to cost efficiency by identifying the patterns and relationships of keywords through associated rule analysis and social network analysis using keywords derived from documents. In addition, it is expected to be used as a decision-making aid to determine the cost efficiency of a specific building construction site by establishing a logistic regression model using core keywords. As a systematic database of construction supervision documents and an integrated system of massive data generated by digital technology are established in the future, the accuracy and reliability of the cost efficiency evaluation model are expected to be reinforced.

키워드 : 공사감리보고서, 성과측정, 비용효율성, 연관규칙분석

Keywords : construction supervisory documents, performance measurement, cost efficiency, association rule analysis

1. 서론

1.1 연구의 목적

건설산업에 디지털 전환이 본격화되면서 디지털 기술의 활용도가 다양해지고 있다. 최근 건축현장 공사감리에 업무의 효율성을 확보하기 위해 디지털 기술 적용에 부단히 노력하고 있다. 공사감리는 건축물의 설계도서대로 시공되는지 품질, 안전, 시공 등의 영역에서 점검하고 지도/감독하는 행위로 정의된다[1]. 민간 건축현장의 성과평가를 위한 체계와 데이터 수집은 여전히 표준화에 한계가 있어 축적된 방대한 데이터를 활용한 성과관리에 어려움이 있다[2,3]. 건축현장 비용효율성 성과측정에 관한 연구는 충분하지 않고, 문서의 디지털화 수준이 아직 낮은 실정이다. 따라서, 본 연구에서 공사감리 문서를 활용해 건축현장의 비용효율성을 평가할 수 있는 방법론과 향후 민간 건축현장에 비용 성과를 측정할 수 있는 모델을 제안하고자 한다.

2. 텍스트마이닝 및 연관규칙 분석 기반 비용효율성 평가

2.1 연구 방법

본 연구는 공사감리에서 생성되는 최종보고서의 종합의견 및 정성적 평가 의견을 활용하여 비용 효율성을 측정할 수 있는 모델을 제안한다. 수집된 43개 현장의 건축공사감리 문서를 대상으로 텍스트마이닝 및 연관규칙분석을 수행하여, 계획 대비 실적 비용 간 비중과 문서에 도출된 긍정적/부정적 키워드와의 관계 분석을 다음과 같이 실시하였다. 첫째, 공사감리 종합의견서에 등장하는 단어의 빈도를 구해 비용 성과와 직·간접적으로 관련 있을 것으로 추정되는 키워드 127개를 선정하였다. 둘째, 한 문장에 등장하는 단어 집합이 하나의 행으로 구성된 데이터셋을 구축하고, 127개 키워드를 제외한 다른 키워드는 제거한 후 연관규칙분석과 소셜네트워크 분석을 수행하였다. 셋째, 연관규칙분석, 소셜네트워크 분석 결과를 활용하여, 11개 핵심 키워드를 선정하였다. 마지막으로, 선정된 핵심 키워드를 독립변수(x)로 두어, 수식 1, 2와 같이 로지스틱 회귀모형을 기반으로 비용효율성(y)을 예측하였다. 이 과정에서, 회귀계수 추정에 L2 규제가 적용된 로지스틱 회귀 방법론을 사용하였다. 또한, 표본을 학습용과 테스트용으로 분리한 후, 학습데이터를 이용해 LOOCV(Leave-One-Out Cross-Validation) 샘플링 방법으로 회귀 계수를 추정하고 그 평균값을 계산하였다. 이렇게 도출된 회귀 모델

1) 한국건설산업연구원, 연구원

2) 한국건설산업연구원, 연구위원, 교신저자(wsyoo@cerik.re.kr)

을 테스트 샘플에 적용하는 것을 100번 반복하여 최종 예측 정확도를 평가하였다. 본 연구는 비용효율성을 계획 공사비 대비 실적 공사비의 비중으로 정의하여, 100% 미만이면 0으로, 100% 이상이면 1로 이진 변환하여 분석에 활용하였다. 수식 2에서 z는 로그 오즈(log-odds) 또는 로짓(logit)으로 불리는 값으로, 로지스틱 회귀계수의 선형결합으로 이루어진다.

$$y_{predict} = \begin{cases} 1 & \text{if } 1/(1+e^{-z}) \geq 0.5 \\ 0 & \text{if } 1/(1+e^{-z}) < 0.5 \end{cases} \quad (1), \quad z = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \beta_5x_5 + \beta_6x_6 + \beta_7x_7 + \beta_8x_8 + \beta_9x_9 + \beta_{10}x_{10} \quad (2)$$

2.2 연구 결과

표 1은 연관규칙분석 결과 중 지지도를 기준으로 상위 13개 규칙을 보인 것으로, 높은 지지도 값을 갖는 규칙 중 ‘관리’ 및 ‘품질’과 연관된 규칙이 빈번하게 생성되었다. 연관규칙분석 결과를 이용하여 그림 1과 같이 소셜네트워크 분석을 수행하였고 ‘관리’, ‘품질’, ‘계획’ 등이 중심성이 높은 것으로 나타났다. 연관규칙분석 및 소셜네트워크 분석 결과를 활용하여 ‘검토’, ‘계획’, ‘공정’, ‘관리’, ‘변경’, ‘사전’, ‘안전’, ‘자재’, ‘점검’, ‘품질’, ‘시공’ 등 11개 핵심 키워드를 선정하였다. 이중 ‘시공’ 키워드를 제외한 10개 키워드의 비중을 이용하여 표 2와 같이 규제 상수(L2)에 따른 로지스틱 회귀모형 기반 비용효율성을 예측하였다. 규제 상수 0.01~0.1에서 예측 정확도가 0.73~0.74로 가장 높게 나타났고, 0.001~0.1에서 낮은 로그로스 값을 보였다.

표 1. 연관규칙분석 결과

규칙 번호	선행사건	후행사건	지지도	신뢰도	향상도
1	품질	관리	0.16	0.66	1.36
2	공정	관리	0.12	0.67	1.38
3	계획	관리	0.12	0.58	1.20
4	안전	관리	0.12	0.68	1.41
5	시공	관리	0.11	0.56	1.15
6	자재	관리	0.09	0.61	1.25
7	계획	품질	0.09	0.41	1.71
8	자재	품질	0.08	0.52	2.15
9	검토	계획	0.07	0.49	2.38
10	시험	품질	0.07	0.79	3.26
11	확인	품질	0.07	0.45	1.84
12	확보	품질	0.06	0.84	3.47
13	시공	품질	0.06	0.32	1.30

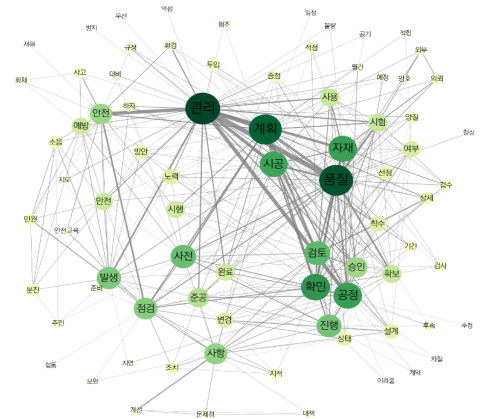


그림 1. 핵심 키워드 간 소셜네트워크 결과

표 2. 규제상수에 따른 이항 로지스틱 회귀모형의 정확도 대비 로그로스

규제 상수(L2)	0.001	0.01	0.1	1	10	100	1000
예측 정확도	0.60	0.73	0.74	0.70	0.68	0.66	0.66
로그로스	0.64	0.61	0.65	0.84	1.25	1.61	1.89

3. 결론

건축현장에서 다양한 디지털 기술 적용이 본격화되면서 자료의 디지털화가 가속되고 있다. 본 연구에서 제시된 비용효율성 평가 모델은 텍스트 기반의 문서 자료를 활용하여, 민간 건축현장에서 생성되는 공사감리문서를 활용하였다. 향후 비용효율성뿐 아니라 일정효율성 평가 모델 구축 후 문서 기반 비용·일정 통합 성과평가 지원 시스템으로 확장될 것으로 기대된다.

감사의 글

본 연구는 국토교통부 디지털 기반 건축시공 및 안전감리 기술개발 사업의 연구비지원(RS-2022-00143493, 과제번호: 1615012983)에 의해 수행되었습니다.

참고문헌

1. 국토교통부. 건축법 및 동법 시행령. <https://www.law.go.kr/건축법>. 2022.
2. 성유경 외. 민간 건축현장 성과측정을 위한 감리보고서 활용성 분석. 한국건축시공학회 학술발표대회 논문집. 2022. pp. 217-218.
3. 이강욱. 건설공사 사후평가 성과분석 체계 개발. 한국산학기술학회 학술발표대회 논문집. 2022. pp. 1000-1001.