

XGBoost를 활용한 교면포장 손상유형 예측

이용준*, 박경훈*, 선종완*

*한국건설기술연구원 구조연구본부/도로관리통합클러스터
e-mail:yongjunlee@kict.re.kr

Prediction of Damage Type of National Highway Bridges Pavement Using XGBoost

Yongjun Lee*, Kyung-Hoon Park*, Jong-Wan Sun*

*Dept. of Infrastructure Safety Research, Korea Institute of Civil Eng. and Building Tech.

요약

본 연구에서는 머신러닝 기법 중 의사결정나무 기법인 XGBoost를 활용하여 일반국도 교면포장에 발생하는 손상유형을 예측할 수 있는 모델을 제안하였다.

1. 서론

도로포장이나 교면포장의 경우 열화나 파손의 예측은 포장 구조 자체의 복잡성과 교통하중, 이상기후 등의 상호 작용을 포함한 여러 가지 이유로 발생하기 때문에 정확하게 예측하는 것은 매우 어렵다. 이에, 관련 기관들은 여러 환경인자들을 고려할 수 있고 장래 도로포장 상태에 따라 전략적으로 대처할 수 있는 예측 모델 개발의 필요성을 강조하였다. 본 연구에서는 일반국도 교량관리시스템(Bridge Management System; BMS)에서 제공하는 교면포장 상태 데이터를 수집하여 머신러닝 기법 중 분류학습 방법인 XGBoost를 활용하여 교면포장 손상유형 예측 기법을 제안하고자 한다.

2. 데이터 수집

분석에 필요한 데이터는 한국건설기술연구원에서 관리하는 BMS에서 수집하였다. 1969년부터 2019년까지 국내에서 건설된 3162개의 교량에 대한 정밀점검 및 정밀진단 6,216건과 정밀점검 및 정밀진단을 통해 수집된 12,972개의 교면포장 검사 결과를 수집하였다. 수집된 데이터에는 16개의 기본정보(지역, 관리기관, 교량 속성 정보 등), 5개의 구조정보(교량 및 포장 형식, 포장 두께 등), 2개의 교통정보(연평균 일교통량, 연평균 일트럭교통량), 16개의 환경정보(기온, 강우, 강설 등), 2개의 점검정보(상태등급, 손상유형)로 구성되어있다. 또한 데이터의 품질 향상을 위해 이상치와 결측치를 제거하였으며, 누락된 값이 포함된 행은 제거하였다.

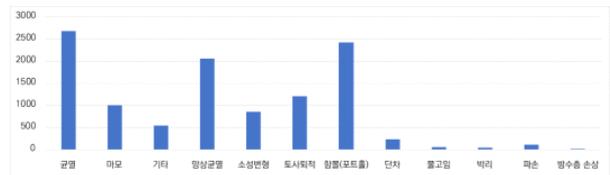
일반 도로포장의 경우 영상차량을 이용하여 정량적 지표

인 균열, 러팅, IRI 등 포장상태를 조사하여 평가하는 반면에 교면포장의 경우 [표 1]과 같이 점검자의 육안점검을 통해 손상유형 및 손상물량을 조사하고 총 4단계(A~D등급)로 평가하고 있다.

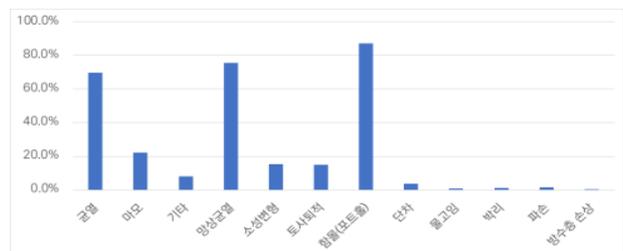
[표 1] 교면포장 손상유형

점검부재	손상종류
아스팔트	균열, 함몰, 단차 및 요철, 블리딩, 마모
콘크리트	균열, 마모, 박리, 파손

수집된 데이터의 교면포장 손상유형별 발생건수는 균열, 함몰(포트홀), 망상균열이 주로 발생하고 있으며, 교면포장의 유지보수가 필요한 상태등급(C등급 이하)에서 주로 발생한 손상유형으로는 함몰(포트홀)이 가장 높은 비율을 차지하였다.



[그림 1] 교면포장 손상 유형

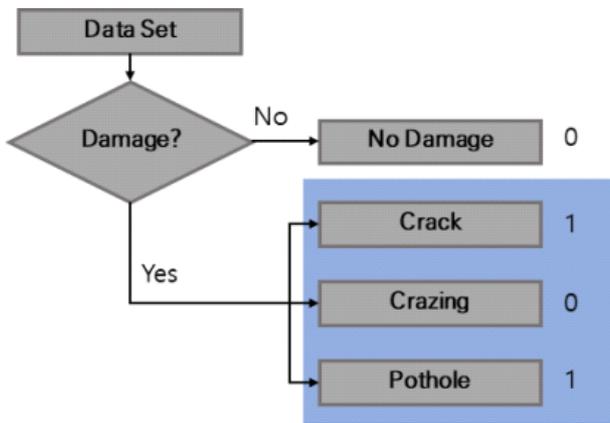


[그림 2] C등급 이하 교면포장 손상유형 비율

이에 본 연구에서는 교면포장에 주로 발생하고 있는 균열, 포트홀, 망상균열과 같은 손상에 대한 발생여부를 예측하는 모델을 개발하고자 한다.

3. 머신러닝을 활용한 파손예측 모델 개발

본 연구에서 개발한 모델은 [그림 3]과 같이 데이터를 투입하면 파손 발생 여부를 분류하고, 파손이 발생한 경우 손상 유형을 예측하도록 설계하였다.



[그림 3] 모델 설계구조

XGBoost 모델의 파라미터 튜닝을 위해서 먼저 Learning Rate(0.1~0.3)을 선택하고 이 학습률에 맞는 Tree 개수를 선정하였다. 그리고 max_depth, min_child_weight, gamma 등의 파라미터를 선정하고 학습률을 조정된 후 위의 과정을 반복하여 GridSearchCV를 통해 최적 파라미터를 선정하였다. 현재 수집되지 않은 미래의 실제 데이터를 투입하는 경우를 모사하기 위해, 데이터를 8:2 비율로 학습(training)과 검증(testing) 데이터로 분할하였으며, 학습 데이터는 다시 5등분하여, 그 중 하나를 학습된 모델의 성능을 확인(validation)하는 데이터로 확인하고 나머지 4/5를 학습에 활용하는 과정을 확인 데이터셋으로 바꾸는 과정을 총 5회 반복하여 수행하는 5-fold cross validation을 사용하여 검증하였다. 분류 예측 모델에 사용되는 성능 평가 지표는 [표 2]와 같이 학습된 분류 모델이 예측을 수행하면서 얼마나 오차가 발생하는지 함께 보여주는 지표인 오차행렬(Confusion Matrix)를 기반으로 산정이 가능하다.

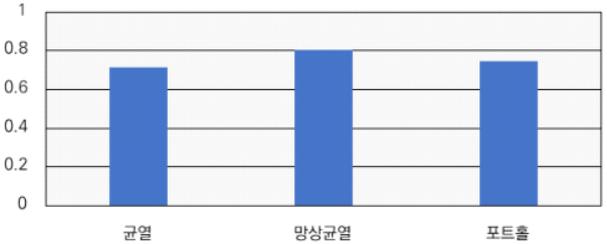
[표 2] Confusion Matrix

Actual \ Predict	True	False
True	True Positive(TP)	False Positive(FP)
False	False Negative(TN)	True Negative(TN)

본 연구에서 성능 지표로는 F1 스코어를 사용하였으며 산정식은 다음과 같다.

$$F1\ Score = 2 / (1 / Precision + 1 / Recall) \quad (1)$$

예측모델의 성능은 F1 스코어 0.7(Kohavi et al, 1997)을 기준으로 하였으며 본 연구에서 개발한 손상유형별 예측모델의 F1 스코어를 산정한 결과 기준치인 0.7을 넘어 성능을 만족한 것으로 나타났다.



[그림 4] F1 Score 결과

4. 결론

본 연구에서는 일반국도 교면포장 손상유형 예측을 위해 머신러닝 기법 중 의사결정나무 기법인 XGBoost 활용하여 교면포장에서 주로 발생하는 손상에 대해 예측하는 방안을 제안하였다.

1. 예측모델의 손상유형별 F1 스코어가 0.72~0.80으로 기준을 만족하였다.
2. 향후 손상유형뿐만 아니라 손상물량까지 예측 하여 유지보수방법 결정할 수 있는 모델을 개발한다면 관리자가 장래 유지보수 계획 수립시 중요한 기초정보로 활용될 수 있을 것이라 기대된다.

참고문헌

[1] <https://nbms.kict.re.kr>, 교량통합관리시스템, 국토교통부.
 [2] 이용준 외, “XGBoost를 활용한 고속도로 콘크리트 포장 파손 예측”, 한국건설관리학회, 제 20권 6호, pp. 34-43, 11월, 2020년.