

ORIGINAL ARTICLE

# 머신러닝 기법과 TBM 시공정보를 활용한 토압식 실드TBM 굴진율 예측 연구

강태호<sup>1</sup>, 최순욱<sup>1\*</sup>, 이철호<sup>1</sup>, 장수호<sup>2</sup>

<sup>1</sup>한국건설기술연구원 지하공간안전연구센터 수석연구원, <sup>2</sup>한국건설기술연구원 건설산업진흥본부 선임연구위원

## A Study on Prediction of EPB shield TBM Advance Rate using Machine Learning Technique and TBM Construction Information

Tae-Ho Kang<sup>1</sup>, Soon-Wook Choi<sup>1\*</sup>, Chulho Lee<sup>1</sup>, and Soo-Ho Chang<sup>2</sup>

<sup>1</sup>Senior Researcher, Underground Space Safety Research Center, Korea Institute of Civil Engineering and Building Technology

<sup>2</sup>Senior Research Fellow, Construction Industry Promotion Department, Korea Institute of Civil Engineering and Building Technology

\*Corresponding author: [soonugi@kict.re.kr](mailto:soonugi@kict.re.kr)

Received: November 26, 2020

Revised: November 30, 2020

Accepted: November 30, 2020

### ABSTRACT

Machine learning has been actively used in the field of automation due to the development and establishment of AI technology. The important thing in utilizing machine learning is that appropriate algorithms exist depending on data characteristics, and it is needed to analysis the datasets for applying machine learning techniques. In this study, advance rate is predicted using geotechnical and machine data of TBM tunnel section passing through the soil ground below the stream. Although there were no problems of application of statistical technology in the linear regression model, the coefficient of determination was 0.76. While, the ensemble model and support vector machine showed the predicted performance of 0.88 or higher. it is indicating that the model suitable for predicting advance rate of the EPB Shield TBM was the support vector machine in the analyzed dataset. As a result, it is judged that the suitability of the prediction model using data including mechanical data and ground information is high. In addition, research is needed to increase the diversity of ground conditions and the amount of data.

**Keywords:** EPB Shield TBM, Machine learning, Algorithm, Geotechnical data, Machine data

### 초록

최근 AI 기술의 발전과 정립으로 자동화 분야에서 머신러닝 기법의 활용이 활발하게 이루어지고 있다. 머신러닝 기법의 활용에 있어 중요한 점은 데이터 특성에 따라 적합한 알고리즘이 존재한다는 점이며, 머신러닝 기법 적용을 위한 데이터세트의 분석이 필요하다. 본 연구에서는 다양한 머신러닝 기법을 기반으로 하천 하부의 토사지반을 통과하는 토압식 실드TBM 터널 구간의 지반정보와 굴진정보를 사용하여 토압식 실드TBM의 굴진율을 예측하였다. 선형회귀모델에서 모델의 통계적인 유의성과 다중공선성에서는 문제가 없었으나 결정계수가 0.76으로 나타났고 앙상블 모델과 서포트 벡터 머신에서는 0.88이상의 예측성능을 보여, 분석한 데이터세트에서 토압식 실드TBM 굴진성능예측에 적합한 모델은 서포트 벡터 머신임을 알 수 있었다. 현재 도출된 결과로 볼 때, 토압식 실드TBM의 기계데이터와 지반정보가 포함된 데이터를 활용한 굴진성능 예측 모델의 적합성은 높다고 판단된다. 추가적으로 지반조건의 다양성과 데이터양을 늘리는 연구가 필요한 것으로 판단된다.

**핵심어:** 토압식 실드TBM, 머신러닝, 알고리즘, 지반정보, 기계데이터



## 1. 서론

건설현장은 열악하고 고위험인 작업환경으로 인하여 작업인력의 고령화가 진행되는 반면, 신규인력의 유입은 감소하고 있다. 인력보충을 위해 해외 인력을 수급하고 있지만, 문화적인 차이와 소통의 문제로 사고위험 및 품질문제가 증가하고 있는 추세이다. 이러한 문제를 해결하기 위해 타 산업분야뿐만 아니라 건설분야에서도 로봇 및 자동화의 적용이 증가하고 있지만, 공장과 같이 반복적이지 않은 작업 공정으로 인하여 한계가 있는 실정이다.

TBM은 터널분야에서 자동화가 적용되기 유리한 장비 중 하나이다. 근래 TBM 자동화에서 관심이 높은 분야는 운전자동화라고 할 수 있다. 말레이시아 MMC-GAMUDA사는 실시간 기계 데이터 모니터링과 AI 알고리즘을 이용한 TBM 장비 제어를 통해 운전자(operator)가 수동으로 개입해야하는 작업을 감소시킴으로써 TBM 작업의 전체적인 효율을 높일 수 있게 하였다. 이 기술은 ‘Autonomous TBM’이라 불리고 있으며, AI알고리즘이 탑재된 모듈을 구성하고 모듈과 PLC(Programmable Logic Controller)를 연결하여 굴진과 스티어링(steering), 슬러리 시스템을 자동화시켰다(Chin, 2020). 그러나 각 시스템에 적용한 AI알고리즘에 대한 상세사항은 알려진 바 없다.

최근 AI(Artificial Intelligence) 기술의 발전과 정립으로 자동화 분야에서 머신러닝(machine learning) 기법의 활용이 활발하게 이루어지고 있다. 머신러닝 학습방법으로는 인공신경네트워크, 서포트 벡터 방법, 뉴로-퍼지방법, 하이브리드 모델 및 입자군집 최적화방법 등이 있다(Gholamnejad and Narges, 2010, Grima et al., 2000, Armaghani et al., 2017, Mahdevari et al., 2014, Salimi et al., 2016, Yagiz et al., 2009, Yagiz and Karahan, 2011). 중요한 점은 데이터 특성에 따라 적합한 알고리즘이 존재한다는 점이며, 머신러닝 기법 적용을 위한 데이터세트(dataset)의 분석이 필요하다.

Kim et al.(2020)은 암반지반에서 TBM 굴진 시 가장 중요한 영향요소를 일축압축강도로 보고 암반구간 이수식(slurry) 쉴드 TBM 현장의 지반조사결과와 기계데이터를 이용하여 데이터세트를 구성하고 규제(regularization)모델과 서포트벡터머신(Support Vector Machine), 앙상블 학습(ensemble learning) 모델을 활용하여 가장 예측 성능이 좋은 스택킹 모델을 암석강도 예측을 위한 최종 모델로 선택한 연구를 진행하였다.

Mokhtari(2020)에 따르면 토압식 쉴드TBM 터널 굴착현장에서 TBM 운전자가 가장 많이 제어하는 항목은 쉴드잭 스트로크(shield jack stroke)와 스크류 컨베이어(screw conveyor) 등이며, 토압식 쉴드TBM 굴진에 영향을 미치는 매개변수로는 추력(thrust), 커터 헤드 토크(torque), 폼 유량(foam flow) 및 스크류 컨베이어 토크 등이 있다고 하였다. 토압식 쉴드TBM에서 이러한 매개변수들은 지반조건에 따라 변화하며 그중에서도 지반 종류에 따라 쉴드TBM 운영 작업이 달라지게 된다.

본 연구에서는 다양한 머신러닝 기법을 기반으로 하천 하부의 토사지반을 통과하는 토압식 쉴드TBM 터널 구간의 지반정보와 굴진정보를 사용하여 쉴드TBM의 굴진율을 예측하였다. TBM터널의 시공정보는 데이터전처리를 통해 데이터세트로 구성하고 머신러닝 기법별 적합한 하이퍼파라미터(hyperparameter)를 선정하여 현재 데이터세트에서의 TBM 굴진율 예측에 적합한 최적모델을 도출하는 연구를 수행하였다.

## 2. 머신러닝기법

### 2.1 선형회귀분석(Linear regression)

선형회귀 분석은 독립변수와 종속변수의 관계를 설명할때 가장 많이 쓰는 회귀분석으로 독립변수(X)와 종속변수(Y) 간의 인과

관계를 찾기 위한 방법이다(식 (1)과 식 (2)). 회귀분석에 사용되는 독립변수 개수에 따라 단순 선형회귀 (simple linear regression), 다중 선형 회귀(multiple linear regression)로 구분된다.

$$Y = \beta_0 + \beta_1 X + \epsilon \quad (1)$$

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \epsilon \quad (2)$$

여기서,  $\beta_0, \beta_1, \dots, \beta_n$  : 가중치,  $\epsilon$ : 오차항

일반적으로, 선형회귀 분석을 사용하기 위해서는 다음의 가정을 만족해야 한다.

- 독립변수와 종속변수는 선형 관계를 가진다.
- 데이터에 아웃라이어 없고, 오차항은 자기 상관성 없다.
- 오차항의 평균은 0이며, 분산은 정규 분포를 가진다.
- 변수간의 상관관계가 존재 하지 않으며, 독립변수와 오차항 서로 독립적이다.

## 2.2 축소방법(Shrinkage Methods)

선형회귀분석에서는 변수가 늘어날 경우 모델이 복잡해지고 회귀계수(가중치)의 팽창으로 인한 특정 학습데이터에 과대적합 (overfitting)이 발생할 수 있는데, 축소방법에서는 이를 억제하기 위해 규제화(regularization)하여 가중치 간 편차를 줄이는 방식으로 과대적합을 방지한다. 대표적인 축소방법으로는 릿지(Ridge), 라쏘(Least Absolute Shrinkage Selector Operator, LASSO), 엘라스틱 넷(Elastic-Net) 3가지의 방법이 있다. 릿지 회귀(Ridge regression, L2 규제)는 가중치를 영에 가깝게 제약하여 과대적합이 되지 않도록 모델을 강제하는 방법이고, 라쏘 회귀(LASSO regression, L1 규제)는 문제가 되는 가중치를 영으로 까지 만들어 완전히 제외하여 과대적합을 방지한다. 그리고, 엘라스틱 넷은 L1, L2 규제를 결합한 방식으로 릿지와 라쏘가 혼합된 방식으로 두 규제의 장단점도 혼합되는 특징을 가지고 있다.

## 2.3 서포트 벡터 머신(Support Vector Machine, SVM)

서포트 벡터 머신은 패턴 인식과 자료 분석 등으로 자주 활용되는 기계학습 모델로서 러시아의 수학자 Vapnik(1995)이 제안하였다. 주어진 샘플에 대해 규칙을 찾아내는 기법으로 주로 분류(classification)나 회귀(regression)의 목적으로 활용된다. 샘플 특성이 정의하는 변수들로 표현된다고 할 때, 샘플들은 차원 데이터 공간에 분포하고, 이 차원의 데이터 공간에서 샘플 그룹들을 구분해하여 최적 경계를 찾아내는 알고리즘이다. 서포트 벡터 머신의 분류는 일정한 마진(margin) 오류 안에서 두 클래스의 폭이 가장 최대가 되도록 하는 반면, 회귀에서는 제한된 마진 오류 안에서 가능한 많은 샘플이 들어가도록 학습하는 방법이다. 이때 경계 내부의 폭을  $\epsilon$ 로 조절한다.

## 2.4 결정트리 회귀(DecisionTree Regression)

1960년대에 소개된 결정트리 기법은 질문을 통하여 데이터 분류를 진행하고 규칙을 찾는 방법으로서 트리 기반으로 한 번 분리

를 진행한 후, 변수를 구분하는 영역을 만들어 원하는 클래스 값을 구한다. 또한, 그룹을 구분하는 분류에서 그룹별로 특징을 발견하거나 어느 집단에 속하는지 여부를 파악하여 세분화 하는데 효율적인 기법이다. 결정트리 기법은 연속형, 범주형 변수에 모두 사용 가능하고 차원축소나 변수 선택 등 의사결정 생성과정을 통해 많은 변수 중에서 상대적으로 종속변수에 주는 영향을 파악할 수 있다(Breiman et al., 1984). 변수의 정규화나 표준화 같은 전처리가 필요하지 않으며, 특정 변수의 값이 누락되어도 사용 가능하나, 학습데이터에 과대 접합되는 특징 때문에 새로운 데이터에 적용하면 예측 성능이 좋지 않아 일반화가 어렵다는 단점이 있다.

## 2.5 앙상블 기법(Ensemble Learning Method)

더 좋은 예측 성능을 위해 다수의 학습 모델, 학습 알고리즘을 따로 쓰지 않고 여러 개의 학습 모델을 동시에 이용해서 최적의 해답을 얻는 방법이다. 앙상블 기법에는 보팅(voting classifier), 배깅(bagging), 부스팅(boosting, bootstrap aggregating), 스택킹(stack) 등이 있다. 보팅은 투표를 통해서 최적 결과를 도출하고, Breiman(1996)이 제안한 배깅은 샘플을 중복으로 생성해서 정답을 찾아가는 방식으로 대표적인 방법으로 랜덤 포레스트(Random forest, Breiman, 2001)가 있다. 부스팅은 여러 개의 모델을 학습하고, 이전 오차를 보완하여 가중치를 부여하고 다시 학습을 진행하는 방식이다. 마지막으로 스택킹은 여러 모델을 기반으로 예측된 결과를 통해 메타(meta) 모델로 재학습하여 다시 결과를 예측하는 방법이다.

## 3. 데이터 분석 및 머신러닝 기법 적용

### 3.1 현장 개요 및 지반 정보

본 연구에서 활용한 토압식 쉘드TBM 시공정보는 노선 인근에 주요 단층이 4개 존재하고 단층의 영향으로 풍화대와 연약파쇄대가 발달한 지역이며, 약 30 m 이상의 연약한 퇴적층이 상부에 형성되어 있다. 하천 하부의 토사지반을 통과하는 TBM 터널 구간에서 약 30개의 시추가 이루어 졌으며, 표준관입시험(Standard Penetration Test, SPT)에서 N값이 1.5~40(심도 15~30 m)으로 나타났다. 터널 심도에서의 평균 N값이 10 이하인 구간은 전체 시추공의 67%이고, N값이 5 이하로 연약한 구간은 37%이다. 터널 구간에는 전반적으로 모래질 점토(CL)와 실트질 모래(SM)가 혼재되어 있으며, 시작부에는 모래질 점토, 실트질 모래가 종점부에는 약간의 실트질 점토(ML)가 분포하는 것으로 나타났다(Fig. 1).

일반적으로 터널설계 시에 기본물성 시험, 수리특성 시험, 연약지반특성 시험, 탄성파 탐사 등의 다양한 지반조사를 수행하고 있지만, 기본적인 시험을 제외하고는 대부분의 시험이 이뤄지는 대상이 일부지역에 국한되어 있다. 터널이 통과하는 구간에 해당하는 각 시험 결과가 대체로 일부만 존재하고 그렇기 때문에 많은 시험 결과들이 분석 자료로 활용하기에 어려움이 있다. 따라서 본 연구에서는 토압식 쉘드TBM 굴진면의 토압 관리에 관련되고 터널 통과구간 전체에서 확보가 가능한 지반특성인 토질분류, N 값, 터널 상부 토피고 등을 굴진을 예측을 위한 분석 자료로 사용하였다.

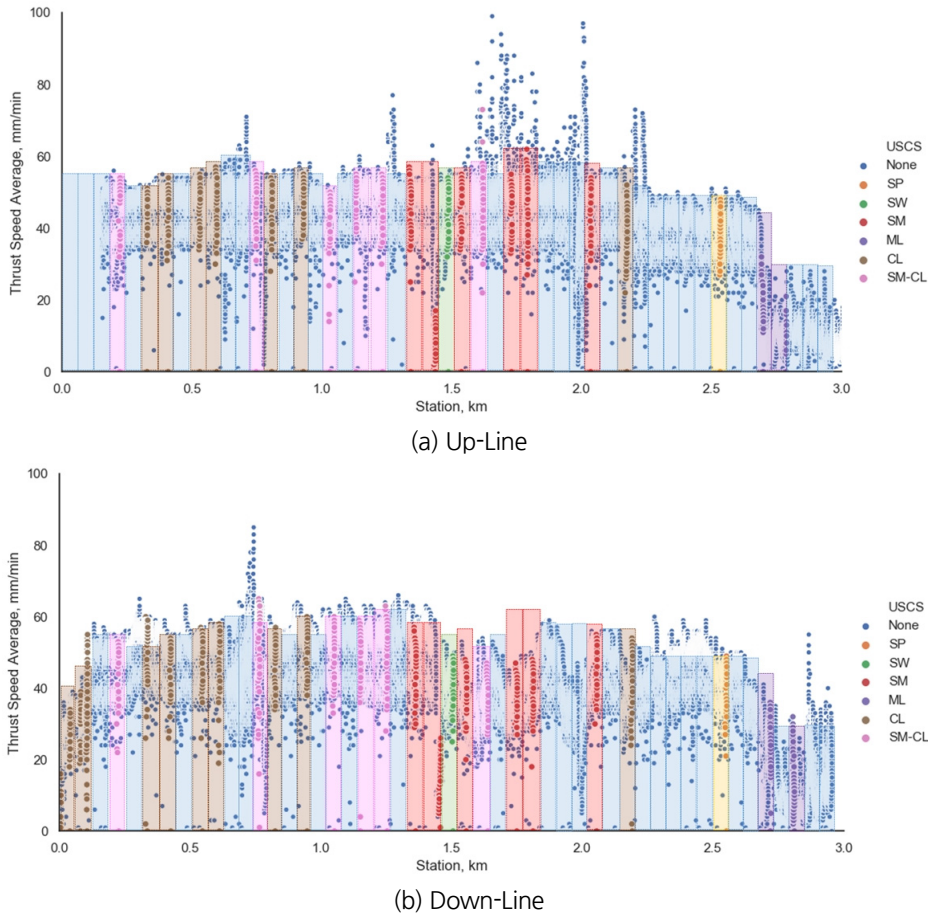


Fig. 1. Geological strata of the TBM tunneling site

### 3.2 굴진자료 및 데이터 준비

본 연구에서 분석한 토압식 쉴드TBM 장비의 주요 제원을 보면, 커터헤드 최대 토크 용량은 9.6 MN·m, 최대 추력은 50 MN 이다(Table 1). 현장에서 사용된 추력은 장비 최대 성능의 약 45.2%이고, 커터헤드 토크는 장비 최대 토크의 56.8% 범위에서 사용된 것을 확인할 수 있었으며, 다양한 토질 조건에 따라 TBM 운전작업이 달라지는 것을 데이터 분포 특성에서 알 수 있다(Fig. 2).

Table 1. Summary of TBM specification

TBM type	EPB
TBM outside diameter (mm)	7,800
Max. shield jack thrust force (MN)	50 (2 × 25 shield jack)
Max. cutterhead torque (MN·m)	9.6
Max. RPM (rev/min)	0.89
Max .Screw revolution (rev/min)	14.2

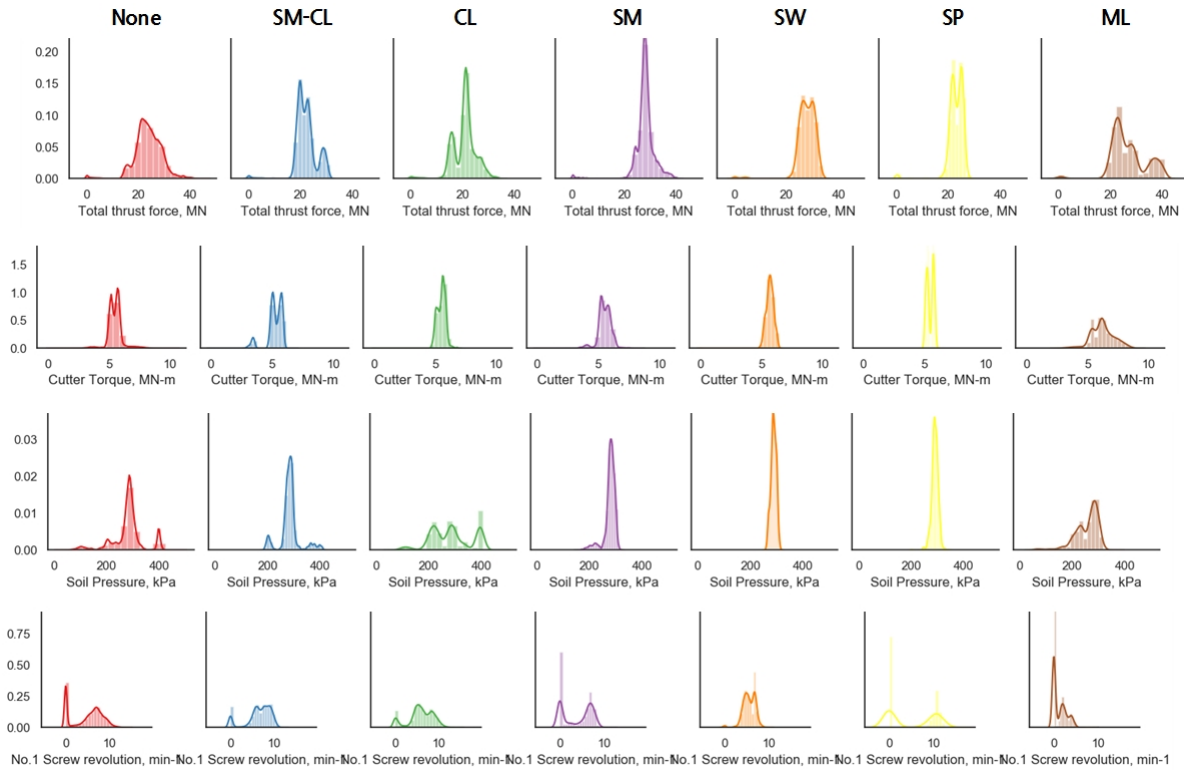


Fig. 2. Distribution of data according to soil conditions for the EPB TBM

본 연구에서는 기계정보와 지반정보가 포함된 데이터를 분석하여 토압식 실드TBM의 굴진성능 예측을 위한 효율적이고 정확한 기계 학습 모델을 구축하고자 하였다. 효과적인 학습 모델을 구축하기 위해서는 데이터의 특성을 정확하게 파악하고 학습에 적합하게 구성하는 것이 필요하다.

실드TBM 굴진 현장에서는 실드 잭 추력, 커터헤드 토크, 스크류 컨베이어 속도, 각종 장비 압력, 온도, 주입제 주입 속도, 위치, 자세 정보, 시간 이력 등을 포함하여 광범위한 데이터 600 여개 이상의 기계 데이터가 기록되고 있다. 이렇게 실시간으로 기록되는 데이터에는 통신 상의 문제나 센싱 등의 문제로 이상치와 결측치가 발생하게 되며, 이러한 이상치와 결측치는 자료의 평균과 표준편차에 영향을 미치고, 최종적으로는 예측성능 저하의 원인이 되기 때문에 TBM 장비 제원과 굴진 데이터를 확인하여 이상치와 결측치는 제거해야한다.

본 연구에서는 데이터 전처리(pre-processing) 과정으로 데이터 값과 평균의 거리를 표준편차로 나눈 작업으로 표준화(standardization)를 적용하고, 결측치 처리는 결측 값이 존재하는 데이터 전체를 제거하는 방식을 사용하였다. 또한 이러한 데이터에서 토압식 실드TBM의 굴진 성능 예측에 영향을 주지 않는 다양한 전압 측정값, 계측 온도, 위치 및 자세 정보와 같은 매개 변수를 분석에서 제외하였다. 최종적으로는 변수 간 상관도 분석을 통하여 기계정보와 지반정보가 포함된 주요 영향 특성 15개를 선정하였다(Fig. 3).

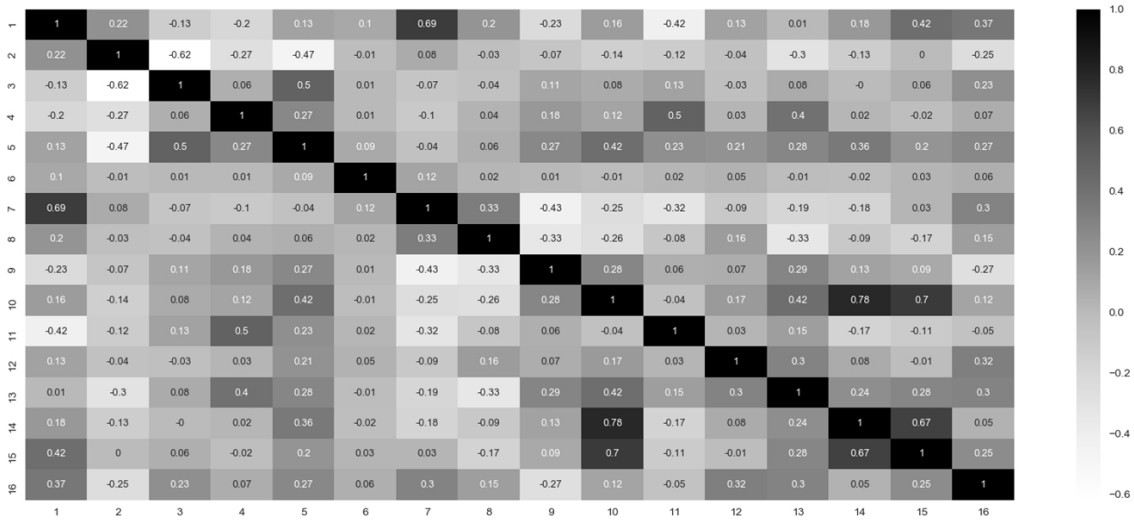


Fig. 3. Pair-wise Correlation between EPB Data features

본 연구에서는 토사 구간 터널 상·하행 굴진의 기계 데이터(Fig. 1) 중에서 시추를 시행한 위치에 지반정보와 매칭이 되는 구간의 데이터 약 16,000개로 분석용 데이터를 구성하였다. 학습 모델 구축을 위한 데이터 분석을 위해 앞서 2절에서 언급한 지도학습 기반 알고리즘을 선정하고, 학습 모델의 적합성을 평가하기 위해서 학습 데이터와 테스트 데이터를 8:2 비율로 나눠서 적용하였다. 선정된 모델의 예측성을 평가하는 방법으로는 오차제곱에 루트를 더하여 실제 관측 값과 차이를 확인할 수 있는 RMSE(Root Mean Squared Error)와 예측 모델이 주어진 자료와 얼마나 적합한가를 평가하는 척도인 결정계수(coefficient of determination)를 사용하였다.

### 4. 분석결과

토압식 쉴드TBM 굴진 성능 예측 모델을 첫 단계로 기본적인 다중선형회귀분석을 수행하였다. 선형회귀분석은 주요 영향 특성인 독립변수와 종속변수 사이의 관계를 확인하고 독립변수의 유의성을 검증하기 위한 작업이다(Fig. 4). 본 연구에서 수행한 선형 회귀 모델의 유의확률(P-value)은 0.0003으로써 통계적 가설 검증 기준인 0.05이하로 이 모델은 통계적으로 유의하다고 할 수 있다. 또한, 독립변수사이의 상관관계를 확인하는 분산팽창요인(Variance Inflation Factors, VIF)검토에서 15개의 독립변수 모두 10이하로 다중공선성 기준에 만족하였다(Table 2). 다중선형회귀 분석 결과에서 RMSE 값은 6.125로 나타났고 결정계수는 0.76으로 선형 회귀 분석의 예측능이 높지 않음을 알 수 있었다(Table 4).

Table 2. Summary of the VIF factors of multiple linear regression model

Features	Total thrust force	Average shield jack stroke	No.1 Screw revolution	Discharge amount of soil	Cutter Torque	No.1 Screw Torque	Cutter Speed	Discharged soil	No.1 Gate stroke	Soil pressure (Down)	Soil Pressure	Thrust Num Selected	N value	USCS	Ground Level
VIF Factors	3.08	1.03	2.82	1.52	1.61	3.93	2.97	1.27	1.71	3.57	3.45	1.58	1.89	2.11	1.77

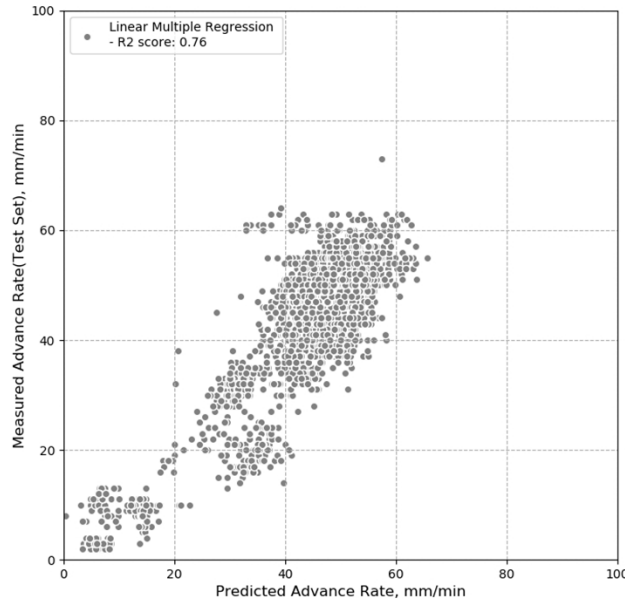


Fig. 4. Results of comparison between Measured and predicted Advance rate for multiple linear regression model

하이퍼파라미터(hyperparameter), 즉 초매개변수는 머신러닝 모델을 만들 때 직접 지정해주는 값이다. 데이터 학습 과정에서 알고리즘 모델에 적합한 파라미터를 조합을 찾기 위한 하이퍼파라미터 튜닝(hyperparameter tuning)이 필요한데, 이는 알고리즘을 모델링 할 때, 모델 성능에 큰 영향을 줄 수 있는 하이퍼파라미터에 대해서 최적에 가까운 값을 찾기 위한 하이퍼파라미터의 조정 작업이라 할 수 있다. 본 연구에서는 모든 조합에 대해서 교차검증(Cross Validation, CV) 후 성능비교를 통해 가장 좋은 성능을 내는 하이퍼파라미터 조합을 찾는 그리드서치(GridSearchCV) 튜닝 기법을 선정하고 7가지 학습모델에 대해 적절한 파라미터를 결정하였다(Table 3).

Table 3. Summary of Hyperparameter for Models

Model	Hyperparameter	
Shrinkage Methods	Ridge	Alpha: 1.0
	LASSO	Alpha: 0.01
	ElasticNet	Alpha : 0.06, L1 ratio: 0.01
Tree-Based Methods	Decision Tree	Max depth : 6
	Bagging - RandomForest	Max depth : 6
	Boosting - GradientBoosting	Learning rate: 0.3, Max depth: 4, Mmax features: 3, n estimators: 45
Support Vector Machine	SVR	C: 10, gamma: 0.1, epsilon:0.1

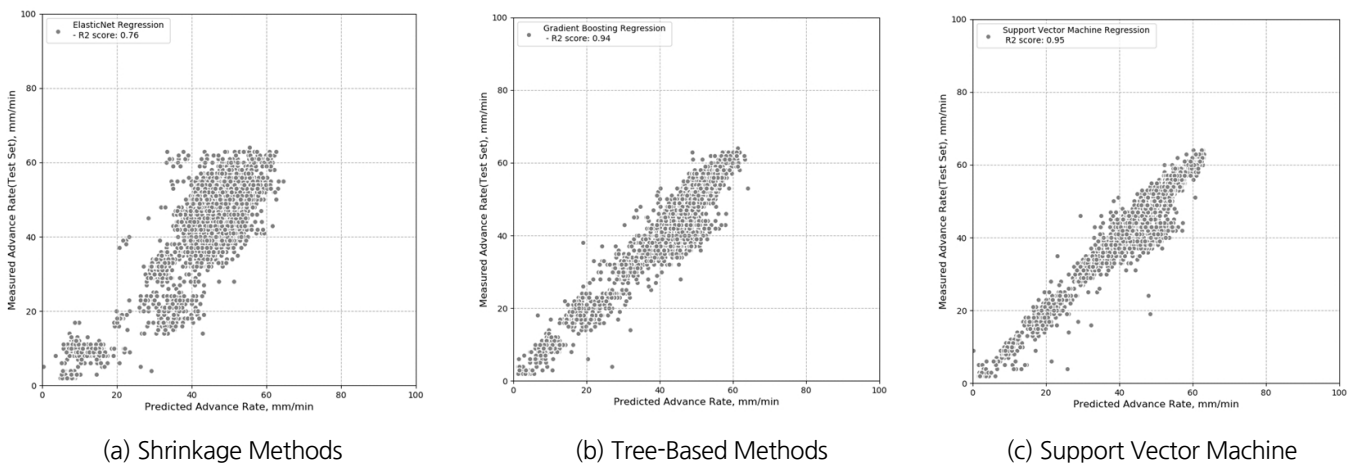
토압식 실드TBM의 굴진성능 예측을 위한 기본 선형회귀 분석과 기계학습 알고리즘 모델을 이용한 분석결과를 Table 4와 Fig. 5에 정리하였다. 모델의 적합성을 나타내는 결정계수는 1에 가까울수록 예측성능이 우수하고 RMSE는 오차를 나타내므로 작을수록 오차가 작아 예측성능이 우수하다는 것을 의미한다. 다중회귀분석과 축소방법인 릿지, 라쏘, 엘라스틱 넷은 결정계수가 0.76 정도이고, RMSE가 6.2 정도로 예측 성능이 나쁘지는 않지만 그렇다고 크게 좋다고도 할 수 없다. 결정트리 기반의 알고리즘에서는



기본 결정트리, 앙상블 모델인 랜덤포레스트, 그랜드 부스팅(GradientBoosting)의 예측 성능이 결정계수 0.88~0.94이고, RMSE 값도 4.3~3.1로 나타나서 부스팅 모델의 예측성능이 높은 것을 확인할 수 있다. 마지막으로 서포트 벡터 머신의 경우 결정계수가 0.95이고 오차인 RMSE 2.75로 가장 작아 예측성능이 좋은 것을 알 수 있다. 최종 분석 결과에서 7가지 예측 모델의 성능을 비교해 보면 앙상블 모델인 그랜드 부스팅과 서포트 벡터 머신이 분석한 데이터셋에 대해서 토압식 쉴드TBM 굴진성능예측에 적합한 모델임을 알 수 있다.

**Table 4.** Results of regression modelling for EPB TBM datasets

Model	Test Dataset		
	R2	RMSE (mm/min)	
Multiple linear	0.760	6.195	
Shrinkage Methods	Ridge	0.766	6.255
	LASSO	0.766	6.255
	ElasticNet	0.762	6.299
	Decision Tree	0.886	4.257
Tree-Based Methods	Bagging - RandomForest	0.884	4.308
	Boosting - GradientBoosting	0.942	3.102
Support Vector Machine	0.952	2.754	



**Fig. 5.** Results of comparison between Measured and predicted Advance rate for regression modelling

## 5. 결론

본 연구에서는 다양한 머신러닝 기법을 기반으로 하천 하부의 토사지반을 통과하는 토압식 쉴드TBM 터널 구간의 지반정보와 굴진정보를 사용하여 토압식 쉴드TBM의 굴진율을 예측하였다. TBM터널의 시공정보는 데이터전처리를 통해 데이터셋으로 구성하고 머신러닝 기법별 적합한 하이퍼파라미터를 선정하여 현재 데이터셋에서의 TBM 굴진율 예측에 적합한 최적모델을 도출하는 연구를 수행하였다.

데이터세트 구성을 위한 지반정보를 구성하는 데 있어, 터널 설계 시 수행된 지반조사 결과는 일부지역에 국한된 자료가 많아서 터널 통과구간 전체에서 확보가 가능한 지반특성인 토질분류, N 값, 터널 상부 토피고 등을 굴진율 예측을 위한 분석 자료로 사용하였다.

변수 간 상관도 분석을 통하여 기계정보와 지반정보가 포함된 주요 영향 특성 15개를 선정하였고 지반정보와 매칭되는 구간의 기계데이터 약 16,000개를 사용하여 데이터세트를 구성하였다. 또한 모델의 검증을 위해 학습 데이터와 테스트 데이터를 8:2 비율로 나눠서 적용하였다.

선형회귀모델에서 모델의 통계적인 유의성과 다중공선성에서는 문제가 없었으나 결정계수가 0.76으로 나타났고 앙상블 모델과 서포트 벡터 머신에서는 0.88이상의 예측성능을 보여, 분석한 데이터세트에서 토압식 쉘드TBM 굴진성능예측에 적합한 모델은 서포트 벡터 머신임을 알 수 있었다. 현재 도출된 결과로 볼 때, 토압식 쉘드TBM의 기계데이터와 지반정보가 포함된 데이터를 활용한 굴진성능 예측 모델의 적합성은 높다고 판단된다. 그러나 이 결과는 본 연구에서 사용한 데이터에 한정될 수 있으므로 추가적으로 지반조건의 다양성과 데이터양을 늘리는 연구가 필요하다.

## 감사의 글

본 연구는 국토교통부 국토교통과학기술진흥원의 스마트건설기술개발사업(과제번호: 20SMIP-A157075-01)인 “교량 및 터널의 원격, 자동화 시공을 위한 핵심기술 개발”의 지원으로 수행되었습니다.

## REFERENCES

- Armaghani, D.J., Mohamad, E.T., Narayanasamy, M.S., Narita, N., Yagiz, S., 2017, Development of hybrid intelligent models for predicting TBM penetration rate in hardrock condition, *Tunn. Undergr. Space Technol.* Vol. 63, pp. 29-43.
- Breiman, L., Friedman, J., Stone, C.J., and Olshen, R.A., 1984, *Classification and Regression Trees*. CRC press.
- Breiman, L., 1996, Bagging Predictors, *Machine Learning*, Vol. 24, No. 2, pp. 123-140. 10.1007/BF00058655.
- Breiman, L., 2001, Random Forests, *Machine Learning*, Vol. 45, No. 1, pp. 5-32. 10.1023/A:1010933404324.
- Chin, J. H. J., 2020, In Pursuit of the Autonomous TBM, *Tunneling Journal*, February/March 2020, pp. 18-22.
- Gholamnejad, J., Narges, T., 2010, Application of artificial neural networks to the prediction of tunnel boring machine penetration rate, *Min. Sci. Technol. (China)* Vol. 20, No.5, pp. 727-733.
- Grima, M.A., Bruines, P.A., Verhoef, P.N.W., 2000, Modeling tunnel boring machine performance by neuro-fuzzy methods, *Tunn. Undergr. Space Technol.* Vol. 15, No.3, pp. 259-269.
- Kim, T. H., Ko, T. Y., Park, Y. S., Kim, T. K., Lee, D. H., 2020, Prediction of Uniaxial Compressive Strength of Rock using Shield TBM Machine Data and Machine Learning Technique, *TUNNEL & UNDERGROUND SPACE*, Vol.30, No.3, pp. 214-225.
- Mahdevari, S., Shahriar, K., Yagiz, S., Shirazi, M.A., 2014, A support vector regression model for predicting tunnel boring machine penetration rates, *Int. J. Rock Mech. Min.*, Vol. 74, pp. 214-229.
- Mokhtari, S., Mooney, M.A., 2020, Predicting EPBM advance rate performance using support vector regression modeling, *Tunn.*

Undergr. Space Technol. Vol.104, 103520.

Salimi, A., Rostami, J., Moormann, C., Delisio, A., 2016, Application of non-linear regression analysis and artificial intelligence algorithms for performance prediction of hard rock TBMs, Tunn. Undergr. Space Technol. Vol.58, pp. 236-246

Vapnik, V.N., 1995, The Nature of Statistical Learning Theory, Springer, New York, pp. 119-166.

Yagiz, S., Gokceoglu, C., Sezer, E., Iplikci, S., 2009, Application of two non-linear prediction tools to the estimation of tunnel boring machine performance, Eng. Appl.Artif. Intell. Vol.22 (4-5), pp. 808-814.

Yagiz, S., Karahan, H., 2011, Prediction of hard rock TBM penetration rate using particle swarm optimization, Int. J. Rock Mech. Min. Sci. Vol.48, No.3, pp. 427-433.