

# Method of Analyzing Important Variables using Machine Learning-based Golf Putting Direction Prediction Model

## 머신러닝 기반 골프 퍼팅 방향 예측 모델을 활용한 중요 변수 분석 방법론

Yeon Ho Kim<sup>1</sup>, Seung Hyun Cho<sup>2</sup>, Hae Ryun Jung<sup>2</sup>, Ki Kwang Lee<sup>3</sup>

<sup>1</sup>Department of Convergence Kinesiology, Kookmin University, Seoul, South Korea

<sup>2</sup>Biomechanics & Sports Engineering Lab., Kookmin University, Seoul, South Korea

<sup>3</sup>College of Physical Education, Kookmin University, Seoul, South Korea

Received : 21 January 2022

Revised : 28 March 2022

Accepted : 29 March 2022

**Objective:** This study proposes a methodology to analyze important variables that have a significant impact on the putting direction prediction using a machine learning-based putting direction prediction model trained with IMU sensor data.

**Method:** Putting data were collected using an IMU sensor measuring 12 variables from 6 adult males in their 20s at K University who had no golf experience. The data was preprocessed so that it could be applied to machine learning, and a model was built using five machine learning algorithms. Finally, by comparing the performance of the built models, the model with the highest performance was selected as the proposed model, and then 12 variables of the IMU sensor were applied one by one to analyze important variables affecting the learning performance.

**Results:** As a result of comparing the performance of five machine learning algorithms (K-NN, Naive Bayes, Decision Tree, Random Forest, and Light GBM), the prediction accuracy of the Light GBM-based prediction model was higher than that of other algorithms. Using the Light GBM algorithm, which had excellent performance, an experiment was performed to rank the importance of variables that affect the direction prediction of the model.

**Conclusion:** Among the five machine learning algorithms, the algorithm that best predicts the putting direction was the Light GBM algorithm. When the model predicted the putting direction, the variable that had the greatest influence was the left-right inclination (Roll).

**Keywords:** Golf, Putting, Artificial intelligence, Machine learning

### Corresponding Author

Ki Kwang Lee

Department of Physical Education,  
University of Kookmin, 77,  
Jeongneung-Ro, Seongbuk-Gu,  
Seoul, 02707, South Korea

Email : kklee@kookmin.ac.kr

## INTRODUCTION

현대 사회에서 골프는 가장 인기 있는 스포츠 중 하나로 남녀를 불문하고 대중화되어 많은 사람들이 즐기는 운동이다(Choi, 2013). 골프가 대중화되며 다양한 골프 장비들이 등장하였으며 골프 자세나 스윙 패턴에 대한 분석 연구들도 많이 진행되었다. 최근에는 스크린 골프나 VR 기반의 골프 프로그램 등이 등장하며 실내에서도 골프를 훈련할 수 있는 시스템들이 많이 등장하고 있다(Nam & Lee, 2017).

특히 스포츠 과학과 스포츠 공학 분야가 발전하며 단순히 경험과 노하우 위주의 훈련 방식으로 선수를 혹사시키던 과거와 달리 과학적인 근거에 의한 신기술 기반의 체계적인 훈련 시스템이 도입되고 있다. 이러한 시도는 선수의 부상을 방지하고 스포츠 능력을 단시간에

성장시킬 수 있도록 도움을 제공한다(Kim, 1995). 특정 분야와 신기술의 결합은 스포츠 분야를 포함한 굉장히 많은 분야에서 우수한 성과를 내며 활발하게 연구되고 있다. 그 중에 인공지능(Artificial Intelligence, AI)은 단기간에 엄청난 성과를 내며 경제, 의료, 생산 등 다양한 분야에 적용 가능한 미래산업의 핵심 기술 중 하나이다.

인공지능이라는 용어는 존 맥카시(John McCarthy)에 의해 처음 사용되었으며(Panesar, 2018), 대량의 데이터를 알고리즘에 학습시켜주면 자동으로 수많은 데이터의 공통된 특성을 빠르게 찾아내어 데이터의 특성에 따라 높은 성능으로 문제를 분류해내는 것이 인공지능 기술의 장점이다. 콘크리트의 균열을 자동으로 분류하는 연구(Seol, Oh & Kim, 2020), 자동차 주행환경에서 보행자를 자동으로 분류하는 연구(Byun & Kwak, 2018) 등 다양한 분야에 적용될 수 있다. 스포츠 분야에서도

최근 머신러닝을 활용한 연구들이 진행되고 있다. 고속으로 이동하는 야구 선수의 위치를 검출하기 위한 방법으로 딥러닝 기반의 CNN 기술을 사용하여 검출하는 시스템을 제안하고 실험 결과를 통해 야구 경기 영상에서 선수의 위치를 효과적으로 검출할 수 있음을 검증한 연구(Kim & Choe, 2019), 골프 스윙 동작을 분석하기 위해 IMU 센서와 머신러닝을 활용하여 분석한 연구(Kim & Park, 2020) 등 스포츠에 머신러닝을 적용하여 분석하려는 다양한 연구들이 시도되고 있다.

이러한 흐름에 따라 본 연구는 골프에서 중요한 동작 중 하나인 퍼팅(Putting)의 방향을 인공지능 기술 중 하나인 머신러닝(Machine Learning)을 적용하여 예측하는 실험을 진행하였다. 골프 전체 동작 중 퍼팅 동작이 55%의 비율을 차지할 정도로 경기의 승패에 있어 퍼팅의 영향력은 엄청난 비중을 차지한다(Lee, 2009). 그러나 대부분의 훈련은 퍼팅보다 스윙 능력을 향상시키는데 집중되어 있다. 골프에서 가장 큰 목표는 아마추어, 프로 선수 할 것 없이 적은 타수로 홀 컵에 공을 넣는 것인데 공을 멀리 보내는 것만 연습해선 안 되고 실수 없이 공을 원하는 곳으로 정확히 보내는 퍼팅 연습도 해야만 한다(Nam & Lee, 2017). Pelz & Frank (2000)의 연구에 따르면 퍼팅 거리에 따라 성공률에 대한 차이가 크게 달라지는 것을 알 수 있다. 프로 선수들을 기준으로 홀과의 거리가 0.91 m 떨어지면 85~90%의 성공률을 보인 반면 3.05 m 떨어진 거리에서는 25% 이상이 볼을 넣지 못했고 4.57 m 이상에서는 10%의 낮은 성공률을 보였다고 보고하였다. 이렇듯 퍼팅을 아마추어가 한 번에 성공시키는 것이 굉장히 어렵기 때문에 오히려 미스 샷이 발생하더라도 최대한 홀과 근접하게 공을 보내는 전략적인 훈련 방식을 연구하여 적용할 필요가 있다(Oh, 2008). 퍼팅은 심리적인 영향을 많이 받기도 하며 일반적인 이론이나 프로 선수들의 퍼팅 방식을 무조건 따라 한다고 실력이 향상되지 않는다. 자신과 맞지 않는 퍼팅 훈련 방식은 오히려 퍼팅 성공에 대한 심리적 불안감을 유발시켜 성공률을 낮추는 원인으로 작용될 수도 있다. 특히 개인이 자신의 퍼팅 스타일이나 패턴을 스스로 분석하고 수정하여 훈련에 적용하는 것은 굉장히 어렵다. 왜냐하면 퍼팅은 스윙에 비해 굉장히 정적이며 움직임이 적어 그 차이를 사람이 일일이 구별해내기 어렵기 때문이다. 또한 전문가가 옆에서 퍼팅에 대한 노하우나 피드백을 제공하여도 개인의 신체 구조적 특징과 습성에 따라 성공률을 높이는 퍼팅 패턴이 다를 수 있어 자신에게 특화된 퍼팅 훈련을 진행하기 힘들다. 선행연구에서는 이러한 퍼팅의 일관성을 향상시키기 위해서는 전문가의 방법을 모방하는 훈련 방식보다 자신만의 퍼팅 자세를 반복 훈련하여 체득해나가는 방식을 추천한다(Jang & Park, 2013).

이러한 필요성에 기인하여 본 연구는 인공지능 기술을 적용하여 퍼팅 방향을 자동으로 예측하는 모델을 구축하였다. 관성을 측정하여 최종적으로 물체가 기울어진 각도를 측정하는 IMU 센서의 골프 퍼팅 데이터를 전 처리하는 프로세스를 구축하고 퍼팅 할 때 홀 근처로 나아가는 최종 방향을 가장 잘 예측하는 머신러닝 알고리즘을 찾기 위해 여러 머신러닝 알고리즘을 구현한 후 각 모델의 성능을 비교하여 최종 제안 모델을 선정하였다. 마지막으로 선정된 제안 모델을 활용하여 학습에 사용된 IMU 센서의 12개 변인 중 방향을 예측하는 데 있어 가장 영향을 미치는 최적의 변인이 무엇인지 찾아내고 방향 예측 모델을 구축할 때 12개의 변인 중 어떤 것이 학습 영향력을 높이는 지 분석하였다. 이러한 모델 구축과 예측 성능에 영향을 미치는 중요 변인을 찾아내는 연구를 진행함으로써 추후 퍼팅 성공률을 높이고 미스

샷을 최소화하도록 도움을 제공하는 트레이닝 시스템, 훈련 장비 개발, 인공지능 기반의 분석 방법론 등의 다양한 연구에 기초 자료로 활용되는 것을 목적으로 삼는다.

## METHOD

본 연구는 골프 퍼팅 경험이 없는 K 대학교 20대 성인 남자 6명을 대상으로 퍼터에 12개의 변인을 측정하는 IMU 센서를 부착하였고 4.5 m 거리에 떨어진 홀을 향해 퍼팅을 실시하는 실험을 진행해 퍼팅 데이터를 수집하였다. 수집된 데이터가 방향 예측 모델의 개발목적에 맞게 빠르고 정확한 학습을 할 수 있도록 전체 데이터의 일관된 패턴을 찾아내는 탐색적 분석(EDA)을 실시하여 분석한 기준에 맞는 데이터 전 처리(Data Preprocessing)를 진행한 후 인공지능 기술인 머신러닝을 활용한 모델링(Modeling)을 실시하여 구축한 데이터 셋을 모델에 학습시켰다. 마지막으로 머신러닝 알고리즘 별 성능을 비교하여 최고의 성능을 나타내는 모델을 제안 모델로 선정한 후 IMU 센서의 12개 변인을 하나씩 적용하여 학습 성능에 영향을 미치는 중요 변인을 분석하였다.

### 1. 참여자

본 연구를 위해 골프 퍼팅 경험이 없는 K 대학교 20대 성인 남자 6명이 참여하였다. 대상자는 3주간 주 2회 하루 50번씩, 6일에 걸쳐 총 300번 퍼팅을 수행하였다. 대상자는 6개월 이내 근골격계 질환이 없는 건강한 상태로 골프 퍼팅을 수행하는데 있어 전혀 문제가 없었으며 실험 전에 퍼팅 절차를 충분히 구두로 설명한 후 실험을 진행하였다. 연구 대상자의 정보는 (Table 1)과 같다.

**Table 1.** Participants' information (Mean  $\pm$  SD)

	N	Age (yrs)	Height (cm)	Weight (kg)
Mean $\pm$ SD	6	24.5 $\pm$ 2.6	174.6 $\pm$ 4.8	69.1 $\pm$ 7.7

### 2. 실험 장비

본 실험에 사용된 장비는 아래 표(Table 2)와 같다.

**Table 2.** Laboratory equipment and data measurement tools

Division	Model name	Manufacturer
Video measurement	Camera	Intel Real Sense
9-DOF sensor	9-DOF sensor	E2BOX
Artificial lawn mat	Artificial Grass mat	Manufactured
Golf putter	Odyssey Red Ball putter	Callaway golf (US)
Golf ball	Diawings M5	Diawings (KR)
Golf hole cup	Silicon hole cup	Top golf mat (KR)

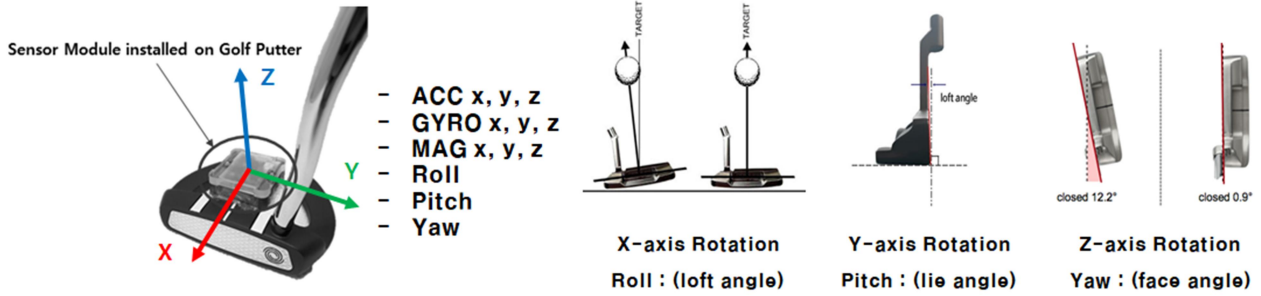


Figure 1. Description of the variables of the IMU sensor

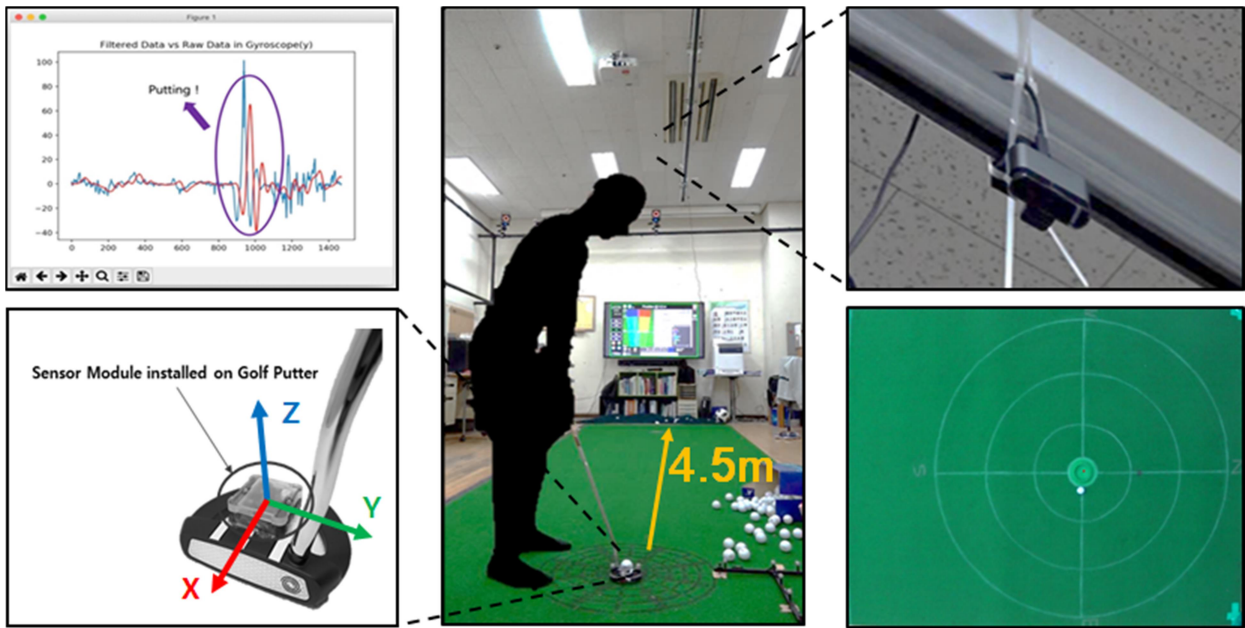


Figure 2. Putting data collection environment.

카메라를 활용해 공의 위치를 파악하고, 관성센서로 퍼터의 신호를 수집하였다(Figure 2). 퍼터 헤드 움직임을 측정하기 위해 9축 IMU 센서(Wireless AHRS EBIMU24GV3)를 활용하였다(가속도  $\pm 16g$ , 각속도  $\pm 2000^\circ/s$ , 지자기  $\pm 4800 \mu T$ ). 각 3축, 총 9가지 데이터를 측정하고 수집된 데이터를 활용하여 퍼팅 동작에 대한 좌우 기울기(Roll), 전후 기울기(pitch), 좌우 흔들기(yaw) 데이터를 산출 후 종합하여 총 12개의 변인으로 구성된 데이터 셋을 구축하였다(Figure 1).

### 3. 실험 방법 및 절차

#### 1) 데이터 수집 및 전처리

퍼팅 데이터 수집하기 위해 퍼팅 시작점으로부터 4.5 m 떨어진 거리에 골프 홀 컵을 고정하였다. 퍼팅 후 골프공의 위치에 따른 레이블링 작업을 위해 홀 라인과 수직으로 약 5 m 위 천장에 카메라를 설치하였다. 성공적인 퍼팅 동작의 시작점을 정의하기 위해 퍼팅 시작마다 3초 카운트를 세는 기계 신호가 끝난 후 연구 대상자가 퍼팅을 실시

하도록 지도하고 3차원 전역좌표계(Global Coordinate System) 설정을 위한 캘리브레이션(Calibration)을 통해 전역 좌표계의 방향은 전후 방향 X축, 좌우 방향 Y축, 수직 방향 Z축으로 설정하였다. 매 퍼팅 동작에 대한 데이터를 퍼터에 부착된 IMU 센서로 측정 후 동영상 데이터를 확인하였다. 이때 골프 홀 컵(20cm) 이내에 위치한 공은 가운데로 판단하고 제외하였으며, 가운데 기준으로 좌측 방향과 우측 방향에 대한 퍼팅 결과를 레이블링 하였다. 또한 (Figure 3)의 데이터 전 처리 프로세스를 활용하여 머신러닝 기반 알고리즘이, 퍼팅 전후를 고려하여 최대한 퍼팅 임팩트 순간을 중점적으로 인식할 수 있도록 퍼팅 전후의 시점과 상관 없는 데이터를 제거하는 방식으로 데이터 셋을 축소시키는 전 처리를 수행하였다.

구축된 최종 데이터 셋은 모델에 학습시키기 위한 훈련용 데이터(Train Data)와 모델의 예측 성능을 검증하기 위한 시험용 데이터(Test Data)로 분할하였다. 그 비율은 머신러닝 분야 데이터 셋 분할에 자주 사용되는 7:3의 비율을 채택하였다. 각 레이블 별 훈련용, 시험용 데이터 셋의 구성은 (Table 3)과 같다.

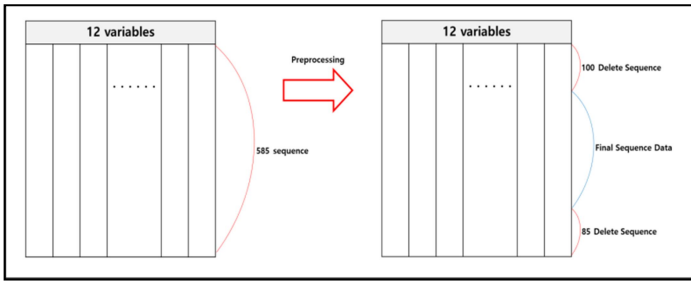


Figure 3. Data preprocessing process

Table 3. Specification of the putting dataset

	Right putting	Left putting	Total
Train	556	556	1,112
Test	185	185	370
Total	741	741	1,482

## 2) 머신러닝 기반 모델 구축 및 중요 변인 분석 방법론

본 연구에서 제안하는 모델 구축과 중요 변인을 분석하기 위한 전체 방법론 구조는 (Figure 4)과 같다.

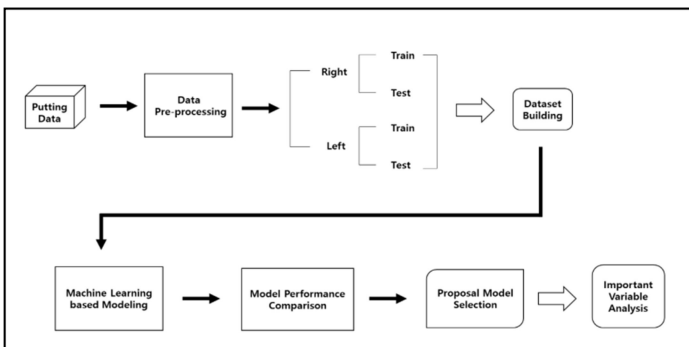


Figure 4. Methodology for building machine learning-based models and analyzing important variables

데이터에 대한 전 처리와 훈련 및 시험용 데이터 셋 분할이 완료되면 훈련용 데이터를 새로운 데이터를 예측할 때 기존 학습되었던 데이터 중 가장 가까운 거리에 위치한 K개 이웃 데이터의 정보를 고려하여 새로운 데이터의 정보를 예측하는 K-NN (K-Nearest Neighbors) 알고리즘(Onel, 2018), 각 데이터의 특성을 개별로 취급해 파라미터를 학습하기 때문에 비교적 고차원의 매개 변수에 민감하지 않아 훈련과 예측 속도가 빠른 Naive Bayes 알고리즘, 문제를 분류하는 결정에 다가가기 위해 '예'나 '아니오' 형식의 질문을 반복하며 학습을 진행하여 가능한 가장 많은 정보를 가진 것을 자동으로 골라내어 시험 데이터에 대한 타깃 값을 예측하는 Decision Tree 알고리즘, 기존 Decision Tree 알고리즘의 주요 단점 중 하나인 훈련 데이터에 대한 과대적합

(Over Fitting) 문제를 방지하기 위해 개발된 Random Forest 알고리즘 (Guido, 2015), 여러 개의 Decision Tree를 묶어 높은 성능의 모델을 구현하는 또 다른 방법의 앙상블 기법인 Light GBM 알고리즘(Guolin et al., 2017) 총 다섯 개의 머신러닝 알고리즘에 각각 집어넣어 학습을 진행하였다. 이때 모델의 성능에 대한 지표는 정확도(Accuracy)를 사용하였다. 정확도는 직관적으로 모델 예측 성능을 나타내는 평가 지표이며 전체 데이터 샘플 중 실제 정답을 예측해낸 비율을 계산하여 모델의 예측 성능을 확인하는 평가 지표로 많이 활용되고 있다.

위와 같은 평가 지표를 활용하여 모델 별 성능을 비교한 후 가장 높은 성능을 보이는 머신러닝 알고리즘을 최종 제안 모델로 선정하고 선정된 제안 모델을 활용하여 퍼팅 시 방향을 예측하는데 있어 중요하게 영향을 미쳤던 변인을 찾아내는 실험을 진행하였다. 이때는 최종 선정된 제안 모델 구축 프로세스에 12개 변인을 각각 하나씩 넣어 모델링을 실시한 후 각 모델들의 성능을 채택한 평가 지표를 사용하여 비교하는 방식으로 중요 변인 분석을 진행하였다. 중요 변인 분석에 대한 방법론 구조는 (Figure 5)와 같다.

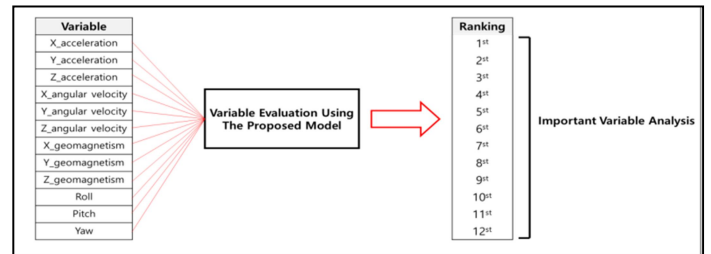


Figure 5. Important variable analysis methodology

## 4. 실험 환경

모델 구축에 실제로 적용된 실험 환경은 (Table 4)와 같다. 하드웨어는 외부 HPC (High Performance Computing) 지원 사업에서 제공받은 Tesla V100 32G GPU를 사용하였으며, 소프트웨어는 Python 3.7.7, Tensorflow 1.15.0, sklearn 0.24.2 등을 실험 환경으로 구성한 후 연구를 진행하였다.

Table 4. Experimental environment

HW	Tesla	V100 32G
SW	Python	3.7.7
	Sklearn	0.24.2

## RESULTS

### 1. 머신러닝 기반 예측 모델 결과

머신러닝 기반 알고리즘 K-NN, Naive Bayes, Decision Tree, Random Forest, Light GBM을 적용하여 퍼팅 예측 모델을 구축한 후 전체 모델

의 훈련 및 시험 정확도를 비교한 결과는 (Figure 6), (Table 5)와 같다.

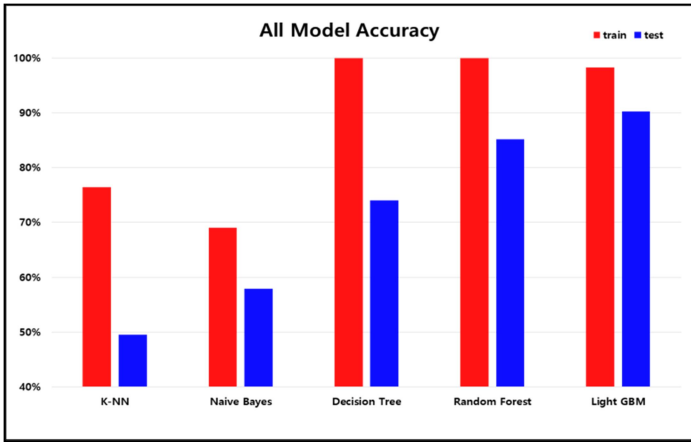


Figure 6. Comparison of training and testing accuracy of the all model

Table 5. Comparison of training accuracy and test accuracy by prediction model

Algorithm	Train_Accuracy	Test_Accuracy
K-NN	76.44%	49.46%
Naive Bayes	69.06%	57.84%
Decision Tree	100%	74.05%
Random Forest	100%	85.14%
Light GBM	98.29%	90.27%

머신러닝 알고리즘인 K-N을 활용하여 구축한 모델의 훈련 정확도는 76.44%로 나타났으며 시험 정확도는 49.46%로 나타났고 Naive Bayes 알고리즘을 활용하여 구축한 모델의 훈련 정확도는 69.06%, 시험 정확도는 57.84%로 나타났다. 그리고 Decision Tree 알고리즘을 활용하여 구축한 모델의 훈련 정확도는 100%, 시험 정확도는 74.05%로 나타났으며 Random Forest 알고리즘을 활용하여 구축한 모델의 훈련 정확도는 100%, 시험 정확도는 85.14%로 나타났다. 마지막으로 Light GBM 알고리즘을 활용하여 구축한 모델의 훈련 정확도는 98.29%, 시험 정확도는 90.27%로 나타났다.

이러한 결과를 토대로 예측 모델 별 성능을 비교하면 다섯 개의 모델 별 훈련 정확도 및 시험 정확도를 비교한 결과 트리 계열의 머신러닝 모델들이 퍼팅 방향 예측에 있어 높은 성능을 나타내는 것을 확인할 수 있었다. K-NN 기반 모델의 경우 훈련 정확도와 시험 정확도의 차이는 26.98%p였으며 Naive Bayes 기반 모델의 경우 훈련 정확도와 시험 정확도의 차이는 11.22%p였다. 또한 Decision Tree 기반 모델의 경우 훈련 정확도와 시험 정확도의 차이는 25.95%p였으며 Random Forest 기반 모델의 훈련 정확도 및 시험 정확도의 차이는 14.86%p를 보였다. 마지막으로 Light GBM 기반 모델의 경우 훈련 정확도와 시험 정확도의 차이가 8.02%p로 다른 네 개의 모델들보다 가장 적은 정확도 차이를 확인할 수 있었다. 훈련 정확도와 시험 정확도의 차이 폭이

크다는 것은 훈련 데이터에 모델이 과대적합이 되었다는 결과이므로 가장 차이 폭이 적은 Light GBM을 최종 퍼팅 방향 예측 제안 모델로 선정하였다.

## 2. 제안 모델의 중요 변인 분석 결과

최종 제안 모델로 선정한 Light GBM 모델의 학습 프로세스를 활용하여 퍼팅 방향을 예측하는데 있어 가장 많은 영향을 미쳤던 중요 변수의 분석을 실시하였다. 총 12개의 변인을 하나씩만 모델 훈련에 활용하여 학습을 완료한 후 모델 별 방향 예측 정확도를 비교하였다. 전체 변인 별 모델의 정확도 비교 결과는 (Figure 7), (Table 6)과 같다.

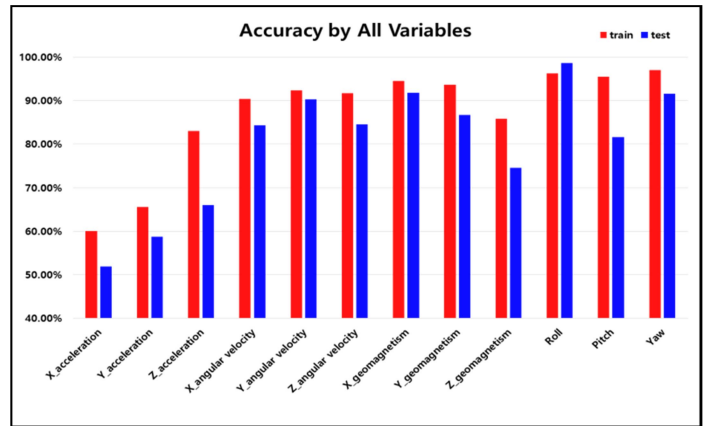


Figure 7. Comparison of accuracy of models by all variables

Table 6. Comparison of training and testing accuracy of Light GBM models by all variables

Variable	Train_Accuracy	Test_Accuracy	Accuracy difference width
X_acceleration	59.98%	51.89%	8.09%
Y_acceleration	65.56%	58.65%	6.91%
Z_acceleration	83%	65.95%	17.05%
X_angular velocity	90.38%	84.32%	6.06%
Y_angular velocity	92.36%	90.27%	2.09%
Z_angular velocity	91.73%	84.59%	7.14%
X_geomagnetism	94.6%	91.89%	2.71%
Y_geomagnetism	93.71%	86.76%	6.95%
Z_geomagnetism	85.88%	74.59%	11.29%
Roll	96.31%	98.65%	-2.34%
Pitch	95.5%	81.62%	13.88%
Yaw	97.03%	91.62%	5.41%

## DISCUSSION

본 연구는 퍼팅 간의 미묘한 차이를 자동으로 분류해낼 수 있는 머신러닝 기술을 활용하여 퍼팅 방향을 예측하는 모델을 구축하고 가장 높은 성능의 모델을 제안 모델로 선정하여 퍼팅의 방향을 예측하는데 있어 중요하게 영향을 미치는 변인을 찾아내는 중요 변인 분석을 실시하였다. 이를 통해 IMU 센서로 측정된 12개 변인만으로도 퍼팅 방향을 예측할 수 있다는 것을 증명하고 높은 성능의 제안 모델을 활용하여 퍼팅 시 나아가는 공의 방향을 예측하는데 중요한 변인을 분석하는 방법론을 제안하였다.

### 1. 머신러닝 기반 예측 모델 별 성능 비교

K-NN, Naive Bayes, Decision Tree, Random Forest, Light GBM 총 다섯 개의 머신러닝 알고리즘을 사용하여 전 처리가 완료된 퍼팅 데이터를 각각 훈련시킨 후 학습이 완료된 모델들의 성능을 비교하였다. K-NN 기반 예측 모델의 경우 훈련 정확도와 시험 정확도가 76.44%, 49.46%로 나타났다. 좌측 및 우측 두 개의 방향을 분류하는 문제에서 시험 정확도가 50% 아래로 떨어진다는 것은 모델이 데이터를 충분히 학습하지 못하는 과소적합(Under Fitting)이 발생한 것으로서 주어진 훈련 데이터만으로 퍼팅 방향을 분류하기에는 너무 단순한 알고리즘이라 데이터에 내재된 구조를 학습하지 못한다는 것을 알 수 있다. Naive Bayes 기반 예측 모델의 경우 훈련 정확도와 시험 정확도가 69.06%, 57.84%로 나타났다. 훈련 정확도는 K-NN 알고리즘보다 성능이 떨어지는 것을 볼 수 있었지만 시험 데이터를 사용한 검증 과정에서 K-NN 기반 모델보다 시험 정확도가 높아진 것을 확인할 수 있다. K-NN의 훈련 정확도 및 시험 정확도의 차이 폭이 26.98%였고 Naive Bayes의 차이 폭이 11.22%였던 것을 고려할 때 K-NN 기반 모델보다 학습에 영향 받지 않은 새로운 데이터를 예측하는 일반화(Generalization)성이 높음을 알 수 있다. Decision Tree 기반 예측 모델의 경우

훈련 정확도와 시험 정확도가 100%, 74.05%로 나타났다. 훈련 정확도와 시험 정확도의 폭이 25.95%나 차이가 나는데 훈련 정확도가 100%라는 것은 굉장히 훈련 데이터에만 과도하게 학습된 과대적합이 발생한 것으로 볼 수 있다. Random Forest 기반 예측 모델의 경우 훈련 정확도와 시험 정확도가 100%, 85.14%로 나타났다. Decision Tree 알고리즘과 같이 훈련 정확도에서 굉장히 높은 결과를 보였으며 시험 정확도는 11.09% 상승한 것을 볼 수 있고 훈련 정확도와 시험 정확도의 차이 폭도 14.86%로 많이 감소된 것을 확인할 수 있다. 이것은 Random Forest 알고리즘이 Decision Tree 알고리즘보다 모델의 일반화성이 높았다는 것을 발견할 수 있다. 이러한 결과를 통해 학습에 사용한 IMU 센서의 데이터는 트리 구조의 알고리즘에서 높은 성능이 나온다는 사실을 발견하였으며 Random Forest와 같이 앙상블을 적용하면 모델의 일반화성이 높아짐을 확인할 수 있었다. Light GBM 기반 예측 모델의 경우 훈련 정확도 98.29%, 시험 정확도 90.27%로 다른 트리 기반 알고리즘보다 훈련 정확도는 낮아졌지만 시험 정확도는 오히려 90%대까지 상승한 것을 볼 수 있다. 훈련 정확도와 시험 정확도의 차이 폭도 8.02%로 굉장히 감소 함으로서 다른 네 개의 알고리즘보다 퍼팅 방향을 예측하는데 있어 예측 성능 및 일반화성이 높다는 것을 확인하였다. 이러한 전체 모델의 성능을 비교해본 결과 퍼팅 방향을 예측하는 최종 제안 모델은 Light GBM 알고리즘이 선정되었다.

### 2. 제안 모델을 활용한 중요 변인 분석

퍼팅 방향 예측 모델의 성능을 비교한 끝에 선정된 최종 제안 모델을 활용하여 모델이 방향을 예측하는데 있어 중요하게 영향을 미쳤던 변인의 중요 순위를 산출하는 실험을 진행하였다. 기존에 가장 높은 성능을 보였던 Light GBM 알고리즘에 변인 12개를 하나씩만 학습을 진행한 후 모델 별 방향 예측 정확도를 비교하였다. 각 변인 별 훈련 정확도와 시험 정확도 그리고 정확도 차이 폭을 고려하여 결정된 최종 제안 순위는 (Table 7)과 같다.

**Table 7.** Ranking of important variables in predicting putting direction

Ranking	Train_Accuracy	Test_Accuracy	Final proposal ranking
1st	Yaw	Roll	Roll
2st	Roll	X_geomagnetism	Yaw
3st	Pitch	Yaw	X_geomagnetism
4st	X_geomagnetism	Y_angular velocity	Y_angular velocity
5st	Y_geomagnetism	Y_geomagnetism	Y_geomagnetism
6st	Y_angular velocity	Z_angular velocity	Z_angular velocity
7st	Z_angular velocity	X_angular velocity	X_angular velocity
8st	X_angular velocity	Pitch	Pitch
9st	Z_geomagnetism	Z_geomagnetism	Z_geomagnetism
10st	Z_acceleration	Z_acceleration	Z_acceleration
11st	Y_acceleration	Y_acceleration	Y_acceleration
12st	X_acceleration	X_acceleration	X_acceleration

제안 모델에서 방향을 예측하는데 가장 크게 영향을 주었던 변인은 좌우 기울기(Roll)였으며 훈련 정확도는 96.31%, 시험 정확도는 98.65%, 정확도 차이 폭은 -2.34%로 나타났다. 훈련 정확도도 높은 수치의 성능을 보였지만 학습에 영향을 받지 않은 데이터로 검증을 실시할 때 오히려 방향을 잘 예측하는 결과를 보여주었다. 이는 좌우 기울기가 모델의 일반화성을 높이는데 기여하는 변인으로 추정되며 전체적으로 높은 성능을 나타내어 1순위의 변인으로 선정하였다.

이러한 기준에 따라 최종 순위는 좌우 기울기, 좌우 흔들기(Yaw), x축 지자기(X\_geomagnetism), y축 각속도(Y\_angular velocity), y축 지자기(Y\_geomagnetism), z축 각속도(Z\_angular velocity), x축 각속도(X\_angular velocity), 전후 기울기(Pitch), z축 지자기(Z\_geomagnetism), z축 가속도(Z\_acceleration), y축 가속도(Y\_acceleration), x축 가속도(X\_acceleration) 순으로 순위가 결정되었다.

구축된 모델이 방향을 결정하는데 있어 가속도는 다른 변인들에 비해 비교적 큰 영향을 미치지 않는 것을 볼 수 있었다. 또한 개별 변수로 퍼팅 예측 시 정확도의 차이 폭이 크거나 예측 성능이 낮다는 것은 전체 변인을 활용한 모델을 구축할 때 모델의 성능을 낮추는 하나의 요인이 될 수 있을 것으로 사료된다. 대부분의 중요 변인이 각도와 연관되어 있다는 결과를 토대로 제안 모델이 퍼팅의 방향을 예측하는데 있어 퍼팅 시 각도와 연관된 변인을 기준으로 결정하였다는 사실을 발견할 수 있었다. 이러한 중요 변인에 대한 분석 결과를 기반으로 산정된 순위는 추후 퍼팅에 관련된 인공지능 기술을 개발할 때 IMU 센서의 여러 변인 중 퍼팅 방향을 예측하는 학습에서 어떤 변인들을 중점적으로 학습시켜야 할지 참고할 수 있는 기초 자료가 될 것으로 기대한다. 또한 본 연구에서 진행된 방법론에 기초하여 추후 신체 분절의 각도나 근육의 활성화 등의 변인을 추가하여 실험을 진행한다면 더욱 견고하고 정확한 퍼팅 방향 예측 모델을 개발할 수 있을 것이다.

## CONCLUSION

본 연구는 골프 경기에서 중요한 기술 중 하나인 퍼팅 동작에 인공지능 기술 중 하나인 머신러닝을 활용하여 퍼팅의 방향을 자동으로 예측하는 실험을 진행하였다. IMU 센서를 퍼터에 부착하여 퍼팅 데이터를 수집하였고 데이터의 클래스 비율을 맞추고 센서에서 측정되는 노이즈를 제거하기 위한 전 처리를 진행하여 데이터 셋을 구축하였다. K-NN, Naive Bayes, Decision Tree, Random Forest, Light GBM 총 다섯 개의 머신러닝 기반 알고리즘을 활용하여 방향 예측 모델을 구축하였고 저유소 휘발유 출하량 예측을 위해 구축된 머신러닝 예측 모델의 성능을 비교하는 연구(Ahn, Jang & Park, 2020), 미세먼지 예측을 위해 구축된 알고리즘의 적합성을 판단하는 연구(Jo, Jung, Kang & Oh, 2019) 등에서 사용한 모델 별 성능 비교 방법을 적용하여 본 연구에서 구축된 모델의 성능을 비교하였다. 인공지능 분야에서 예측 모델의 우수성은 정확도를 활용하여 비교를 실시하고 있기 때문에 모델의 성능 비교를 위해 예측 정확도를 산출하였다. 이러한 과정을 통해 퍼팅 방향을 예측하는 데 있어 가장 우수한 성능을 나타낸 Light GBM 알고리즘을 최종 제안 모델로 선정하였으며 제안 모델의 알고리즘에 12개 변인을 하나씩 집어넣어 변인 별 성능을 비교 분석하였다. 전체 실험 결과를 바탕으로 얻어진 결론은 다음과 같다.

1. 다양한 머신러닝 알고리즘 중 퍼팅의 방향을 가장 잘 예측하는 알고리즘은 Light GBM이다.
2. 머신러닝 기반의 모델이 퍼팅 방향을 예측하는데 있어 가장 크게 영향을 주는 변인은 좌우 기울기이며 각도와 관련된 변인들이 좌측과 우측 방향을 예측하는 모델의 의사결정에 많은 영향을 미친다.
3. 가속도와 관련된 변인들은 모델이 퍼팅의 방향을 예측하는데 중요하게 작용되지 않으며 오히려 성능 저하의 원인이 될 수 있다.

인공지능은 다양한 분야의 문제를 해결하는 핵심 기술로서 단시간에 많은 성과를 내며 기술의 성능을 인정받고 있다. 인간의 힘으로 분류하거나 예측할 수 없는 다양한 문제들을 자동화하고 새로운 해결법을 발견해낼 수 있다는 점에서 인공지능 기술의 적용은 중요한 연구 영역에 하나가 되어가고 있다. 이러한 추세에 따라 스포츠 분야에서도 머신러닝과 딥러닝 등의 인공지능 기술을 적극 활용하는 연구적 시도가 필요하다고 판단되며 현재 스포츠 분야에서도 인공지능을 적용하여 문제를 해결한 연구 결과가 많이 보고된다.

본 연구는 골프 퍼팅의 방향 예측에 머신러닝 기술을 적용하였지만 실외에서 이루어지는 골프 퍼팅 상황을 충분히 반영하여 측정된 데이터가 아니었다는 점에서 연구적 한계가 있었다. 또한 좌측과 우측 방향 외에도 퍼팅 시 발생하는 여러 요인들이 모델 구축에 반영되지 않았기 때문에 추후 연구에서는 신체 분절의 각도, 퍼팅에 사용되는 근육의 활성화 등 다양한 변인을 추가적으로 적용해 봐야 할 필요가 있다. 그러나 IMU 센서에서 측정되는 12개 변인만을 활용해도 퍼팅의 방향을 상당히 높은 확률로 예측할 수 있었다는 점에서 본 연구의 방법론에 의거하여 연구를 확장해나간다면 더욱 우수한 성능으로 퍼팅 상황을 일관되게 예측하는 모델 및 시스템을 개발할 수 있을 것이다. 또한 본 연구에서 활용한 중요 변인 분석 방법론은 추후 다양한 측정 장비를 통해 얻어지는 수많은 변인들이 퍼팅 예측 모델에 얼마나 영향을 미치는지 확인하는 분석 방법론으로서 추후 퍼팅 훈련 장비나 분석 시스템에 적용되는 예측 센서에도 활용될 수 있을 것으로 기대한다.

## ACKNOWLEDGEMENT

This work was supported by the Sports Promoting Fund of Korea Sports Promotion Foundation (KSPO) from the Ministry of Culture, Sports and Tourism.

## REFERENCES

- Ahn, S. J., Jang, D. S. & Park, S. S. (2020). Comparing Gasoline Shipment Prediction Model Performance of Oil Reservoir using Machine Learning. *KIIT*, 18(11), 17-23.
- Byun, Y. H. & Kwak, G. C. (2018). A Transfer Learning and Performance Comparison of Deep Learning Models for Pedestrian Classification under Automobile Driving Environment. *Journal of Advanced Information Technology and Convergence*, 16(10), 83-92.
- Choi, S. G. (2013). Actual golf satisfaction according to screen golf user's participation level. Master's Thesis. Graduate School of Korea

- National University of Education.
- Guolin, K., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q. & Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. Part of Advances in Neural Information Processing Systems 30.
- Guido, S. (2015). Introduction to Machine Learning With Python. O'Reilly & Associates Inc, ISBN9781449369415.
- Jang, J. H. & Park, D. S. (2013). The Concept and Academic Value of Golf. *Korean Society of Golf Studies*, 7(2), 88-97.
- Jo, K. W., Jung, Y. J., Kang, C. G. & Oh, C. H. (2019). Conformity Assessment of Machine Learning Algorithm for Particulate Matter Prediction. *Journal of the Korea Institute of Information and Communication Engineering*, 23(1), 20-26.
- Kim, J. H. (1995). A Study on the Consciousness of High School Golfers. Master's Thesis. Graduate School of Dankook University.
- Kim, J. & Choe, T. (2019), Fast baseball player location detection system using convolutional neural networks for real time broadcast. *KIISE Transactions on Computing Practices*, 25(3), 171-178.
- Kim, M. S. & Park, S. Y. (2020). Golf swing segmentation from a single IMU using Machine learning. *Sensors*, 20(16), 4466.
- Lee, J. W. (2009). Golf mechanics and mechanics golf. Cheongmungag.
- Nam, G. J. & Lee, J. H. (2017). The Effect of feed-back on 7 iron swing according to 3 stance types. *Journal of The Korean Society of Sports Science*, 28(5), 1161-1172.
- Oh, J. H. (2008). Golf Lesson-A Study on the Compensation and Psychological Responses to the Shot After a Missed Shot. Ace Golf Column.
- Onel, H. (2018). Machine learning basics with the K-nearest neighbors algorithm.
- Panesar, A. (2018). Machine Learning and AI for Healthcare (Big Data for Improved Health Outcomes). Apress, 35-37.
- Pelz, D. & Frank, J. A. (2000). Dave pelz's Putting bible : The complete guide to mastering the green. NY: Doubleday.
- Seol, D. H., Oh, J. H. & Kim, H. J. (2020). Comparison of Deep Learning-based CNN Models for Crack Detection. *Architectural Research*, 36(3), 113-120.