

# 결정 트리 앙상블 방법 및 다층신경망을 이용한 유도전동기의 고장진단

왕동훈\*<sup>1</sup>, 이종현\*<sup>2</sup>, 김시진\*<sup>3</sup>, 김민찬\*<sup>4</sup>, 이인수\*\*

## Fault Diagnosis of Induction Motor using an Ensemble Method of Decision Tree and Multilayer Neural Network

Dong-Hun Wang\*<sup>1</sup>, Jong-Hyun Lee\*<sup>2</sup>, Si-Jin Kim\*<sup>3</sup>, Min-Chan Kim\*<sup>4</sup>, and In-Soo Lee\*\*

### 요약

산업 시스템의 발전으로 인해 제조업, 중공업 등에서의 유도전동기 사용이 증가하고 있다. 전동기의 장기간 사용으로 결함이 발생할 수가 있으며 그러한 결함들은 기계의 성능과 수명을 단축시키고 산업적인 측면에서 물질적, 시간적 손해 및 사용자의 안전사고로 이어질 수가 있다. 그러므로 전동기의 고장을 정확하게 진단하는 것은 매우 중요한 문제이다. 본 논문에서는 유도전동기 시뮬레이터에서 정상, 회전자 고장 및 베어링 고장 데이터를 획득하였고 학습으로는 상태별 데이터를 각각 620쌍, 총 1860쌍을 사용하였고 테스트는 학습과 동일한 개수의 데이터를 사용하였다. 유도전동기의 각 상태별 특징을 학습하기 위해 결정 트리 앙상블 방법인 랜덤 포레스트 방법을 사용하였다. 제안한 알고리즘인 랜덤 포레스트와 그래디언트 부스팅 및 다층신경망을 사용하여 성능을 비교하였고, 실험결과 제안한 모델의 분류 성능이 가장 우수함을 확인하였다.

### Abstract

Due to the development of the industrial system, the use of induction motors in manufacturing and heavy industries is increasing. Long-term use of motors can lead to defects, which can shorten the performance and life of the machine and lead to material and time damage and user safety accidents in industrial terms. Therefore, accurate diagnosis of motor failure is a very important issue. In this paper, normal, rotor fault and bearing fault data were obtained from an induction motor simulator, 620 pairs for each state and a total of 1860 pairs of data were utilized for learning, and the same number of data were used for the test. Random forest method, a decision tree ensemble method, was used to learn the characteristics of each state of the induction motor. The performance was compared using random forest, the proposed algorithm, gradient boosting, and multilayer neural network, and the classification performance of the proposed model was confirmed to be the best as a result of the experiment.

### Keywords

induction motor, fault diagnosis, gradient boosting, random forest, multi-layer neural network

\* 경북대학교 전자전기공학부

- ORCID<sup>1</sup>: <http://orcid.org/0000-0002-5180-5373>

- ORCID<sup>2</sup>: <http://orcid.org/0000-0002-9943-4698>

- ORCID<sup>3</sup>: <http://orcid.org/0000-0002-5064-6120>

- ORCID<sup>4</sup>: <http://orcid.org/0000-0003-0714-646X>

\*\* 경북대학교 전자전기공학부(교신저자)

- ORCID : <http://orcid.org/0000-0001-9773-4710>

· Received: Dec. 09, 2021, Revised: Dec. 28, 2021, Accepted: Dec. 31, 2021

· Corresponding Author: In-Soo Lee

School of Electronic and Electrical Engineering, Kyungpook National University,  
80 Daehakro, Bukgu, Daegu, 702-701, Korea

Tel.: +82-53-950-7843, Email: [insoolee@knu.ac.kr](mailto:insoolee@knu.ac.kr)

## 1. 서 론

산업 시스템이 점점 복잡해지고 정교해짐에 따라 유도전동기의 사용이 중공업, 발전소 등 여러 산업체에서 핵심적인 생산 장비로 사용되고 있다. 유도전동기는 내구성이 우수하고 제조가 비교적 간단하여 산업용과 가전용에서 90%, 50% 이상의 점유율을 차지하고 있다[1]. 그러나 장기간 기계를 사용하게 되면 변형이 일어나거나 기계 부품들의 마모가 발생하고 기계 시스템에 복잡한 진동을 유발하며 이는 기계의 성능과 수명을 단축하게 하는 주요 원인이 된다[2]. 만약 유도전동기에 고장이 발생할 경우 경미한 결함이라도 산업적인 측면에서는 큰 물질적, 시간적 손해 및 사용자의 안전사고로 이어지고 나아가 시스템의 전체적인 신뢰도가 낮아진다. 따라서 전동기의 고장을 초기에 진단하여 시스템의 이상 상태를 실시간으로 감지하고 이를 통해 적절한 조치를 취함으로써 유지보수 비용 절감과 높은 안정성 및 신뢰성을 높일 수 있는 알고리즘에 대한 연구가 필요하다.

유도 전동기는 최근 생산 현장에서 주된 동력원 중 하나로 자리 잡고 있으며 이에 대한 고장 진단 연구의 필요성이 커지고 있다. 전동기의 일반적인 결함은 크게 분류하면 기계적 결함과 전기적 결함으로 나눌 수 있다[3]. 기계적인 결함의 경우는 회전 축의 불일치, 고정의 헐거움, 회전자 마찰, 베어링의 결함 등으로 볼 수 있다. 전기적 결함으로는 회전자 절연 등을 의미한다. 유도전동기의 운전 상태를 점검하는 지표로는 온도, 진동, 음향 신호, 전류 신호 등이 있다. 이 중에 진동 신호를 측정하여 유도전동기의 상태를 진단하는 방법이 많이 사용되고 있다[4][5].

최근 이러한 고장진단에 퍼지 로직, SVM(Support Vector Machine), 인공신경망, 의사결정트리 등의 머신러닝 기반 알고리즘 등에 의한 실시간 검출 진단 기법에 대한 연구가 증가되고 있다. 그 중 인공신경망(Artificial Neural Network)은 생물학적 모델의 뇌 구조를 모방하여 설계된 신경망으로 패턴 인식, 식별 및 분류와 같은 다양한 분야에서 사용되고 있으며 입력 및 출력의 관계를 효율적으로 학습할 수

있다. 또한 인공신경망을 사용한 방법은 비선형적 모델 예측에도 좋은 성능을 보인다[6]-[9].

본 논문에서는 유도 전동기의 기계적인 결함 중 가장 발생 빈도가 높은 회전자 고장과 베어링 고장을 진단할 수 있는 시뮬레이터를 제작하였다. 시뮬레이터를 통해 회전자 고장과 베어링 고장에 대한 데이터를 획득한 후 데이터를 머신러닝 기법 중 하나인 다층신경망(Multi-layer neural network)과 의사결정트리의 앙상블 기법인 랜덤 포레스트(Random forest)와 그래디언트 부스팅 알고리즘을 사용하여 고장진단을 수행하고 각 모델의 진단 결과를 비교하여 모델들의 성능을 검증한다.

본 논문의 구성은 다음과 같이 구성되어 있다. 2장에서는 유도전동기의 정상, 회전자 고장, 베어링 고장의 진동 데이터를 획득할 수 있는 시뮬레이터에 대해 기술하였다. 3장에서는 획득한 데이터를 바탕으로 본 연구에서 고장진단을 위해 사용한 분류 모델들에 대해 기술하였다. 4장에서는 제안한 분류 모델들의 학습 방법과 분류 결과에 관해 기술하였다. 마지막으로 5장에서는 본 연구의 결론과 추후 연구의 계획을 기술하였다.

## II. 유도전동기 고장진단 시뮬레이터

본 논문에서는 유도전동기에서 발생하는 고장 중 대표적인 회전자 고장, 베어링 고장의 데이터를 획득하기 위한 시뮬레이터는 그림 1과 같다. 시뮬레이터는 3개의 유도전동기, 진동 센서, DAQ 시스템으로 구성되어 있으며 시뮬레이터에 사용된 유도전동기는 3개로서 동일한 사양이다. 그림 1에서 1번은 정상적인 동작을 수행하는 유도전동기이며, 2번과 3번은 각각 회전자 고장 및 베어링 고장을 유발한 유도전동기이다.

그림 2는 시뮬레이터에서 발생시킨 고장의 유형을 나타내며, 그림 2(a)는 회전자 고장이며, 모터 회전자에 구멍을 뚫어 고장을 유발하였다. 그림 2(b)는 베어링 고장으로서, 작은 금속 파편들을 베어링에 맞물리게 고장을 유발하였다[10].

시뮬레이터를 통해 정상, 회전자 고장, 베어링 고장 상태의 데이터를 획득하였고, 획득한 1개의 데이터 샘플은 그림 3과 같다.

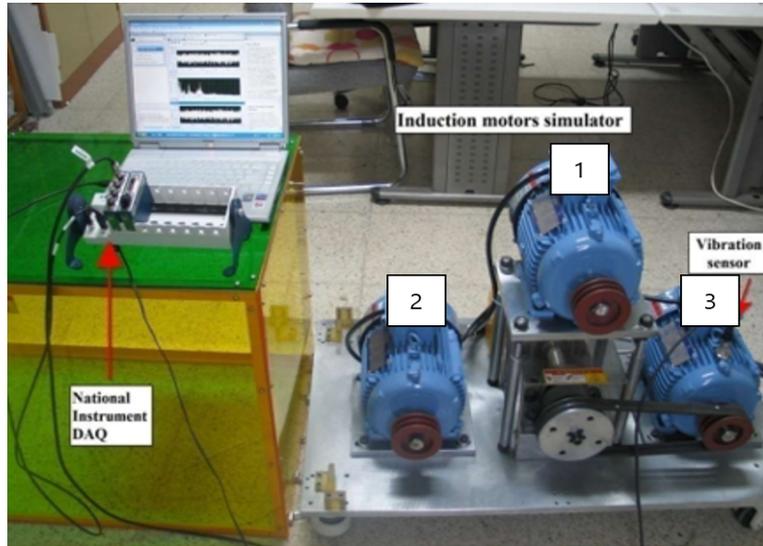
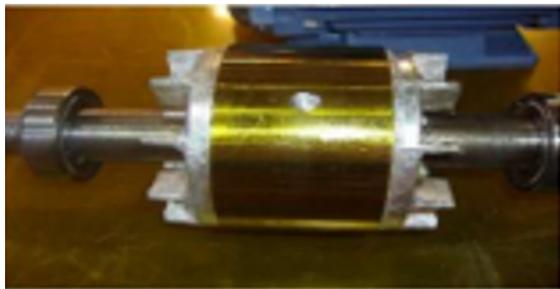


그림 1. 유도전동기 고장진단 시뮬레이터  
Fig. 1. Induction motor failure diagnosis simulator

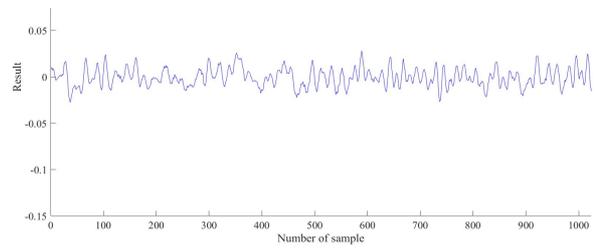


(a) 회전자 고장 모터  
(a) Rotator fault motor

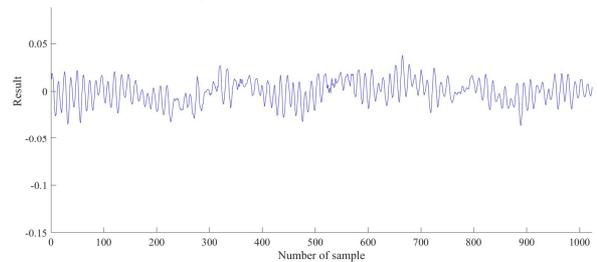


(b) 베어링 고장 모터  
(b) Bearing fault motor  
그림 2. 유도전동기 고장 유형  
Fig. 2. Induction motor fault type

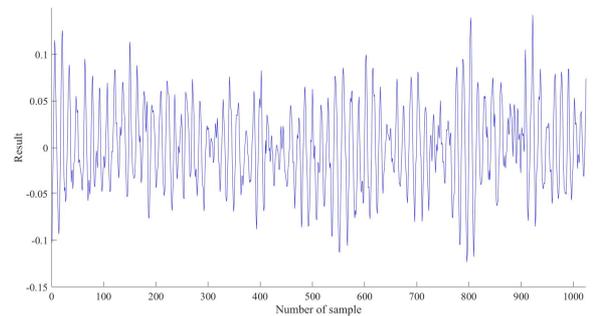
모델 학습과 정확도 테스트를 위한 고장 유형별 데이터 세트는 총 620개로서, 각 데이터 세트에는 1,024개의 데이터 샘플이 포함되어 있다. 정상, 회전자 고장, 베어링 고장에 대한 고장진단을 수행하였고, 그림 3(a)는 정상 데이터, 그림 3(b)는 회전자 고장 데이터, 그림 3(c)는 베어링 고장 데이터이다.



(a) 정상 데이터  
(a) Data of normal state



(b) 회전자 고장 데이터  
(b) Data of rotor failure



(c) 베어링 고장 데이터  
(c) Data of bearing failure  
그림 3. 시뮬레이터를 통해 획득한 데이터  
Fig. 3. Data acquired by simulator

### III. 유도전동기 고장진단 알고리즘

본 논문에서는 다층신경망, 랜덤 포레스트, 그래디언트 부스팅(Gradient boosting)을 사용하여 유도전동기의 고장 진단을 수행하고 성능을 비교하였다. 그림 4는 본 논문에서 제안한 고장 진단 알고리즘을 나타낸다.

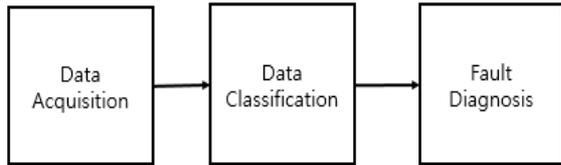


그림 4. 고장진단 알고리즘 다이어그램  
Fig. 4. Schematic diagram of fault diagnosis algorithm

본 논문에서 제안한 고장 진단 시스템은 그림 4와 같이 유도전동기에서 발생한 진동 신호를 디지털화하고 저장하는 데이터 수집 단계, 고장 유형을 분류하는 데이터 분류 단계로 구성되어 있다. 유도전동기에서 얻은 진동 신호 데이터를 분류기를 이용하여 정상, 회전자 고장, 베어링 고장의 세 가지 유형으로 분류한다.

#### 3.1 의사결정 트리

기계 학습에서 의사결정트리(Decision tree)는 어떤 데이터에 대한 일련의 분류 규칙을 통해 데이터를 분류 및 회귀하는 문제에 널리 사용하는 모델이다[11]. 순차적인 질문을 통해 정답을 선택하는 방식으로 학습을 진행한다. 그림 5는 의사결정트리의 구조를 나타낸 것이다.

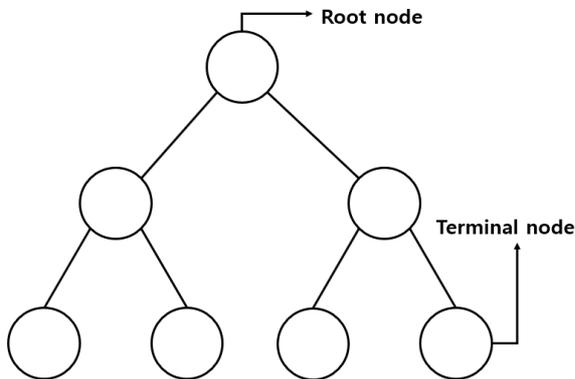


그림 5. 의사결정트리의 구조  
Fig. 5. Structure of decision tree

가장 첫 질문을 하는 분류 기준을 뿌리 노드(Root node)라고 하며, 마지막 질문을 하는 노드를 최종 노드(Terminal node)라고 한다. 결정 트리의 주요 단점은 성능의 변동 폭이 크고, 과적합 되는 경향이 있으며 학습 데이터에 따라 생성되는 결정 트리가 무작위성을 보이며 매우 다르기 때문에 일반화하여 사용하기에 어려움이 따른다. 이와 같은 단점을 보완하기 위해 여러 개의 결정트리들을 사용하는 앙상블 방법을 사용한다.

#### 3.1.1 그래디언트 부스팅

그래디언트 부스팅 알고리즘은 다수의 의사결정 트리를 사용하는 앙상블 방법으로 회귀 분석 및 분류를 수행할 수 있는 모델이며 앙상블 방법 중 부스팅 계열에 속한다. 앙상블 학습은 여러 개의 결정 트리를 결합하여 하나의 결정 트리보다 더 좋은 성능을 내는 머신러닝 기법이다. 앙상블 학습의 핵심은 여러 개의 약 분류기를 결합하여 강 분류기를 만드는 것으로 정확도를 향상시킨다[12]. 그래디언트 부스팅 알고리즘은 오차를 미분한 그래디언트를 줄으로써 모델을 보완하는 방법을 사용한다. 그래디언트 부스팅의 학습 과정은 크게 5단계로 나타낼 수 있다. 우선 초기값에 상수함수를 적용한다. 그 후 2단계는 손실함수를 최소화하는 그래디언트를 구한다. 3단계는 그래디언트를 기본 모델의 타겟 값을 사용하여 학습을 진행한다. 4단계로 학습률을 더해 최종 모형을 만든다. 마지막으로 2 ~ 4단계를 반복한다. 식 (1) ~ (3)은 그래디언트 부스팅의 학습 과정을 나타낸 식이다.

$$F_0(x) = \arg \min_y \sum_{i=1}^n L(y_i, y) \quad (1)$$

$$r_{im} = - \left| \frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right| \quad (2)$$

$$F_m(x) = F_{m-1}(x) + r_m h_m(x) \quad (3)$$

여기서  $\arg \min_y \sum_{i=1}^n$ 는 학습 데이터 세트,  $L(y_i, y)$ 는 손실함수,  $F_0(x)$ 는 초기 값,  $r_{im}$ 은 그래디언트,

$h_m$ 은 트리 모델,  $r_m$ 은 학습률,  $F_m(x)$ 는 출력 값이다. 그림 6은 그래디언트 부스팅의 순차적으로 그래디언트를 업데이트하여 정확도가 높은 분류기를 만드는 학습 구조를 나타낸 것이다.

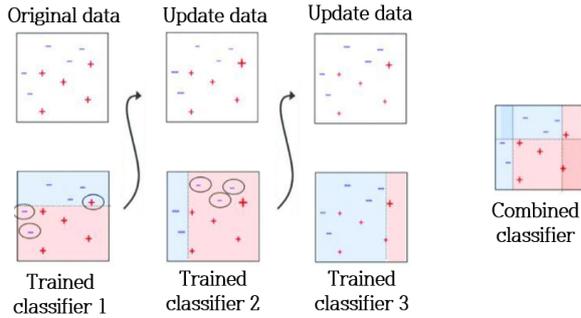


그림 6. 그래디언트 부스팅의 학습 구조  
Fig. 6. Learning structure of gradient boosting

### 3.1.2 랜덤 포레스트

랜덤 포레스트 알고리즘 또한 그래디언트 부스팅 알고리즘과 같이 다수의 의사결정트리를 사용하는 앙상블 방법 중 하나이다. 그래디언트 부스팅 알고리즘과 다른 점은 배깅(Bagging) 방식을 사용한다는 것이다. 배깅 방식은 기존의 여러개의 분류 알고리즘을 사용한 후 투표를 통해 결과를 출력하는 방식과 달리 하나의 분류 알고리즘만을 사용하는 방법이다. 랜덤 포레스트 알고리즘은 다수의 의사결정트리들이 병렬 구조로 학습을 진행하며 이들은 독립적인 개체로 서로 간의 학습에 간섭을 하지 않는다.

각 결정트리들의 학습이 끝나고 최종 노드(Terminal node)에 나온 결과들을 평균치로 예측하거나 과반수 투표를 통해 회귀 분석 및 분류를 수행함으로써 정확도를 높인다[13]. 그림 7은 랜덤 포레스트의 학습 구조를 나타낸 것이다.

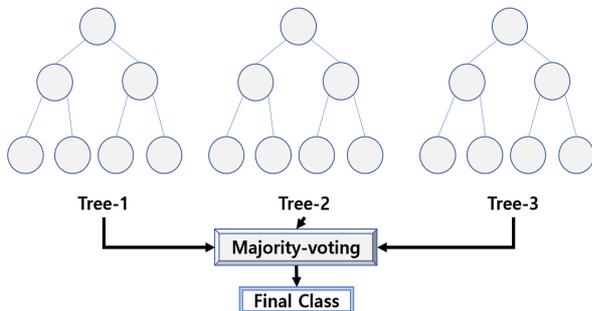


그림 7. 랜덤 포레스트의 구조  
Fig. 7. Structure of random forest

### 3.2 다층신경망

다층신경망은 입력층과 출력층 사이에 하나 이상의 은닉층으로 구성된 구조이며, 그림 8과 같다. 다층 신경망은 순전파 학습과 역전파 학습 방법을 같이 사용하는 신경망으로, 순전파 학습으로 이루어진 오차를 역전파 학습 방식을 사용해서 가중치를 수정함으로써 결과 값을 획득한다[14].

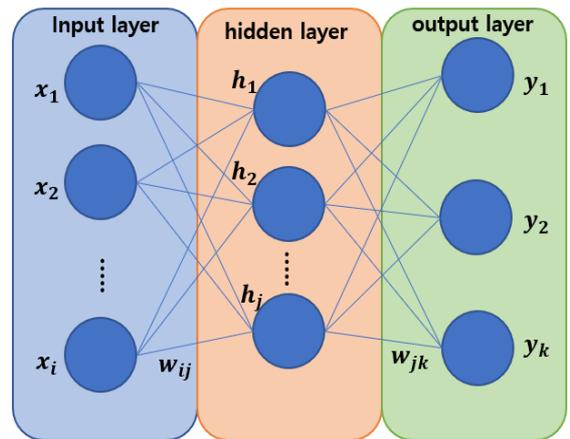


그림 8. 다층신경망의 구조  
Fig. 8. Structure of multi-layer neural network

역전파 학습 방법은 출력층에서 은닉층을 거쳐 입력층까지 역방향으로 가중치를 업데이트해서 오차를 줄이는 방법이다. 식 (4)부터 식 (8)까지 역전파 학습의 수식을 나타내었다. 먼저 식 (4)와 식 (5)을 통해 순전파 학습의 입력값과 출력값을 구할 수 있다.

$$h_j = \sum_i w_{ij} O_i \tag{4}$$

$$O_j = f(h_j + \theta_j) \tag{5}$$

여기서  $h_j$ 는 은닉층의 노드  $j$ 의 입력값,  $w_{ij}$ 는 은닉층의  $j$ 번째 노드에 들어오는 입력층의 가중치,  $f$ 는 활성화 함수이다.  $O_i, O_j$ 는 입력층의 출력값과 은닉층의 출력값이며,  $\theta$ 는 입력 바이어스 값이다. 가중치를 업데이트할 때 필요한  $\delta$ 값은 다음 식으로 구할 수 있다.

$$\delta_i = \lambda(l_i - x_i)f'(O_i) \quad (6)$$

$$\delta_j = \lambda \sum_i \delta_i w_{ji} f'(O_j) \quad (7)$$

여기서  $l$ 은 라벨 값,  $O_i, O_j$ 는 출력값이다. 위 식에서 구한  $\delta$ 를 사용하여 가중치를 감소한다.

$$w_{ji}(t+1) = w_{ji}(t) + \eta \delta_j y_i \quad (8)$$

여기서  $t$ 는 에포크의 수,  $\eta$ 는 학습률이다.

#### IV. 시뮬레이션 결과 및 고찰

본 논문에서는 시뮬레이터에서 획득한 유도전동기의 정상, 회전자 고장, 베어링 고장 데이터를 그래디언트 부스팅, 랜덤 포레스트, 다층신경망 분류 모델의 입력으로 사용하여 학습하였고, 테스트를 통해 고장진단을 수행하였다. 고장 유형별 630개의 데이터 세트를 이용하여 학습하였다. 1개의 데이터 세트는 1,024개의 데이터 샘플로 구성되어 있다.

먼저 그래디언트 부스팅 알고리즘을 이용하여 고장진단을 수행하였다. 의사결정 트리의 수를 150개, 500개, 600개, 700개로 하여 총 4개의 모델로 분류를 진행하였다. 분류 정확도(Accuracy)를 구하기 위한 식은 다음과 같다.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

$$Error = 1 - Accuracy \quad (10)$$

Accuracy는 모든 진단에서 올바른 진단을 한 횟수를 의미하며, Error는 모든 진단에서 잘못된 진단을 한 횟수를 의미한다. 여기서 TP(True Positive)는 정상 데이터를 분류 모델에서 정상으로 진단한 것을 의미하고 TN(True Negative)은 2가지 고장 데이터를 분류 모델에서 각 고장데이터로 진단한 것을 의미한다. FP(False Positive)는 고장 데이터를 분류 모델에서 정상으로 진단한 경우, FN(False Negative)은 2가지 고장 데이터를 분류 모델에서 각

고장)은 정상 데이터를 분류 모델에서 고장 데이터로 진단한 경우를 의미한다[15].

표 1은 그래디언트 부스팅의 분류 정확도 및 분류 오차를 의사결정 트리 개수에 따라 나타낸 것이다. 그래디언트 부스팅 모델들 중에서 의사결정트리 개수가 150개인 모델의 정확도가 97.58%로 가장 낮았다. 의사결정 트리가 500개, 600개, 700개인 경우의 분류 정확도는 각각 99.52%, 99.41%, 99.46%로, 의사결정트리가 500개인 그래디언트 부스팅 모델의 정확도가 가장 높다. 의사결정트리가 600개인 그래디언트 부스팅 모델의 정확도가 낮아진 것을 통해 의사결정트리가 500개인 모델이 최적의 성능임을 확인하였다.

표 1. 그래디언트 부스팅의 분류 정확도 및 오분류율  
Table 1. Classification accuracy and error of gradient boosting

Number of trees	Accuracy (%)	Error ratio (%)
150	97.58	2.42
500	99.52	0.48
600	99.41	0.59
700	99.46	0.54

표 2는 랜덤 포레스트의 분류 정확도 및 오차를 의사결정 트리 개수에 따라 나타낸 것이다. 랜덤 포레스트 모델 중에서 의사결정 트리가 150개인 모델이 분류 정확도가 98.92%로 가장 낮았다. 의사결정 트리가 500개, 600개, 700개인 모델의 분류 정확도는 99.73%, 99.84%, 99.78%이며, 의사결정트리가 600개인 모델의 분류 정확도가 가장 높았다. 의사결정트리가 700개인 모델의 분류 정확도가 낮아진 것을 통해 의사결정 트리가 600개인 모델이 최적의 성능임을 확인하였다.

표 2. 랜덤포레스트의 분류 정확도 및 오분류율  
Table 2. Classification accuracy and error of random forest

Numbers of trees	Accuracy (%)	Error ratio (%)
150	98.92	1.08
500	99.73	0.27
600	99.84	0.16
700	99.78	0.22

마지막으로 다층신경망을 이용한 고장진단을 수행하였다. 본 논문에서 사용한 다층신경망의 은닉층은 총 2층으로 첫 번째 은닉층의 노드 수는 128개, 두 번째 은닉층의 노드 수는 64개를 사용하였고, 두 은닉층의 활성화 함수로 ReLu 함수를 사용하였다. 사용한 epoch의 수는 250개로 다층 신경망 학습을 진행하였다.

표 3은 본 논문에서 사용한 다층신경망의 분류 정확도 및 분류 오차를 나타낸 것이다.

표 3. 다층신경망의 분류 정확도 및 오분류율  
Table 3. Classification accuracy and error of MNN

Method	Accuracy (%)	Error ratio (%)
MNN	98.92	1.08

표 4는 최적의 그래디언트 부스팅 모델, 최적의 랜덤 포레스트 모델, 다층신경망 모델의 분류 정확도를 비교한 것이다. 3가지 분류 모델의 정확도를 비교한 결과 랜덤 포레스트는 99.84%, 그래디언트 부스팅은 99.52%, 다층신경망은 98.92%로, 다층신경망의 분류 정확도가 가장 낮았고 랜덤 포레스트의 분류 정확도가 가장 높았다. 그림 9, 10 및 11은 각각 그래디언트 부스팅, 랜덤 포레스트 및 다층신경망의 분류 결과를 그래프로 나타낸 것이다.

표 4. 사용한 방법들의 분류 정확도와 오분류율  
Table 4. Error and classification accuracy of using method

Methods	Accuracy (%)	Error ratio (%)
Gradient boosting	99.52	0.48
Random forest	99.84	0.16
MNN	98.92	1.08

그래프의 x축은 샘플의 데이터수를 의미하며, y 축은 클래스 분류 값이며 0, 1, 2의 값은 각각 정상, 회전자 고장 그리고 베어링 고장을 의미한다. 파란색 계열은 기존의 라벨 값이며 빨간색 계열은 학습을 통해 분류된 값이다. 그래프를 통해 분류 성능을 비교한 결과 다층신경망을 사용한 분류 정확도는 98.92%, 랜덤 포레스트는 99.84%, 그래디언트 부스팅은 99.52%로 다층신경망이 오차가 큰 것을 확인하였다. 최종적으로 의사결정트리의 랜덤 포레스트 알고리즘의 분류 성능은 그래디언트 부스팅 및 다층신경망에 비해 우수한 것을 확인하였다.

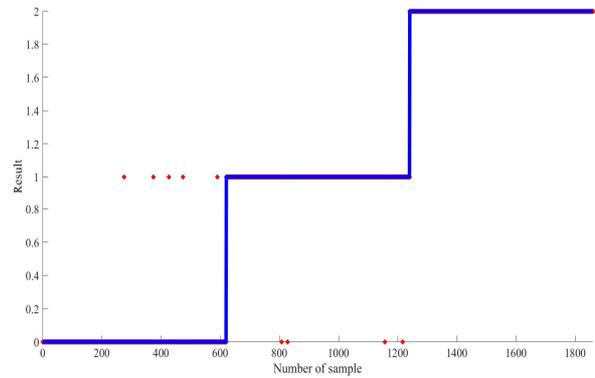


그림 9. 그래디언트 부스팅의 분류 결과  
Fig. 9. Classification result of gradient boosting

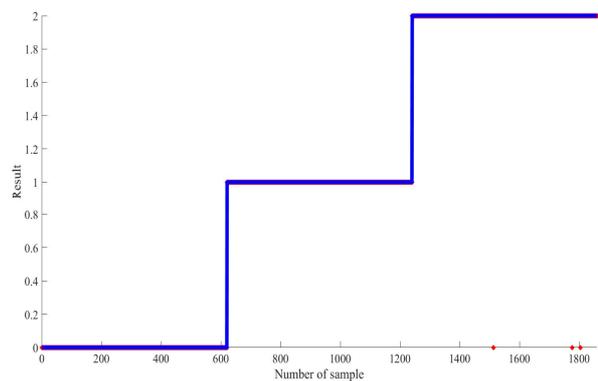


그림 10. 랜덤 포레스트의 분류 결과  
Fig. 10. Classification result of random forest

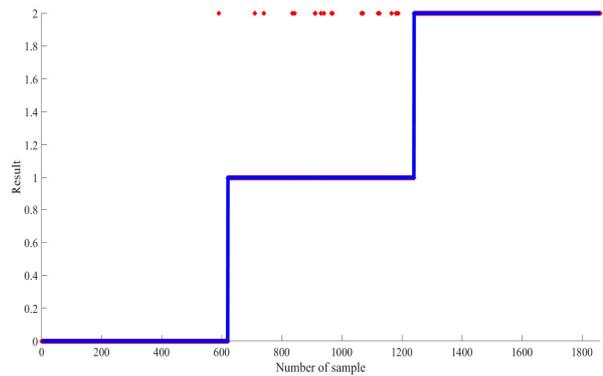


그림 11. 다층신경망의 분류 결과  
Fig. 11. Classification result of multi-layer neural network

## V. 결 론

본 논문에서는 분류기로 그래디언트 부스팅, 랜덤 포레스트, 다층신경망을 사용한다. 유도전동기의 고장 진단을 위해 고장 분류 시뮬레이터를 제작하였고 시뮬레이터 실험을 통해 정상, 회전자 고장,

베어링 고장의 3가지 유형의 데이터를 획득, 제안한 분류기를 통해 고장진단을 수행하였다.

그래디언트 부스팅 및 랜덤 포레스트는 의사결정 트리의 수를 150, 500, 600, 700개로 하여 각각 4개의 분류 모델을 구성하였다. 그리고 다층신경망은 2개의 은닉층에 ReLU 함수를 사용하였고 출력층에 softmax 함수를 사용하였으며 총 250번의 epoch로 학습을 진행하였다. 그래디언트 부스팅 모델에서 의사결정트리가 150개인 모델의 정확도는 97.58%, 500개는 99.52%, 600개는 99.41%, 700개는 99.46%였다. 의사결정트리의 개수가 500개인 모델의 분류 정확도가 가장 높았고 600개인 모델의 분류 정확도는 500개에 비해 낮았다. 랜덤 포레스트 알고리즘의 경우 의사결정트리가 600개인 모델의 정확도가 99.84%로 가장 높았고, 700개인 모델의 분류 정확도는 600개에 비해 낮았다. 따라서 그래디언트 부스팅은 의사결정트리가 500개인 모델이 최고의 성능을 보였고, 랜덤 포레스트는 의사결정트리가 600개인 모델이 최고의 성능을 보였다. 다층신경망 모델의 분류 정확도는 98.92%로 그래디언트 부스팅 및 랜덤 포레스트 모델에 비해 낮았다. 최종적으로 랜덤 포레스트의 분류 성능이 그래디언트 부스팅 및 다층신경망에 비해 우수한 것을 확인하였다.

추후 연구에서는 본 논문에서 수행한 고장진단 모델을 실제 상황에 적용해 고장진단 성능을 비교, 분석하고자 한다.

## References

- [1] Jang, K. B., "Technology Trends to High Efficiency Motors : Focused on Induction Motor", KIEE, Vol. 65, No. 6, pp. 995-1000, Jun. 2016. <http://dx.doi.org/10.5370/KIEE.2016.65.6.995>.
- [2] S. M. Lee, "Fault Diagnosis for Rotating Machinery with Clearance Using HHT", Master's Thesis of Univ. of Sungkyunkwan, 2007.
- [3] P. Vas, Parameter Estimation, "Condition Monitoring, and Diagnosis of Electrical Machines", Clarendon Press, Oxford, 1993.
- [4] I. Y. Onel, and M. E. H. Benbouzid, "Induction Motor Bearing Failure Detection and Diagnosis: Park and Concordia Transform Approaches Comparative Study", IEEE/ASME Transactions on Mechatronics, Vol. 13, No. 2, pp. 257-262, Apr. 2008. <https://doi.org/10.1109/TMECH.2008.918535>.
- [5] I. S. Lee, "Fault Diagnosis System Development of Induction Motors Using Discrete Wavelet Transform and Neural Network", Journal of KIIT, 2011.
- [6] T. H. Cho, H. R. Hwang, J. H. Lee, and I. S. Lee, "Comparison of Intelligent Methods of SOC Estimation for Battery of Photovoltaic System", Advanced Computer Science and Applications, Vol. 9, No. 9, pp. 49-56, Sep. 2018. <http://dx.doi.org/10.14569/IJACSA.2018.090907>.
- [7] N. Saravanan and K. I. Ramachandran, "Incipient gear box fault diagnosis using discrete wavelet transform (DWT) for feature extraction and classification using artificial neural network (ANN)", Expert Systems with Applications, Vol. 37, No. 6, pp. 4168-4181, Jun. 2010. <https://doi.org/10.1016/j.eswa.2009.11.006>.
- [8] E. Byvatov, U. Fechner, J. Sadowski, and G. Schneider, "Comparison of Support Vector Machine and Artificial Neural Network Systems for Drug/Nondrug Classification", Journal of the American Chemical Society, Vol. 43, No. 6, pp. 1882-1889, Nov. 2003. <http://dx.doi.org/10.1021/ci0341161>.
- [9] B. Dębska, and B. Guzowska-Świder, "Application of artificial neural network in food classification", Analytica Chimica Acta, Vol. 705, No. 1-2, pp. 283-291, Oct. 2011. <https://doi.org/10.1016/j.aca.2011.06.033>.
- [10] Mohammed Abdallah and Omar Elkeelany. "A survey on data acquisition systems DAQ", International Conference on Computing, Engineering and Information. IEEE, Fullerton, CA, USA, pp. 240-243, Apr. 2009. <https://doi.org/10.1109/ICC.2009.24>.
- [11] Y. K. Kim and C. S. Hong, "Re-anonymization

Technique for Dynamic Data Using Decision Tree Based Machine Learning", Journal of Korea Information Science Society, Vol. 44, No. 1, pp. 21-26, Jan. 2017. <http://dx.doi.org/10.5626/JOK.2017.44.1.21>.

- [12] H. S. Han, S. J. Cho, and U. P. Chong, "Neural-network-based FAULT Detection and Diagnosis Method Using EIV(errors-in variables)", The Korean Society for Noise and Vibration Engineering, Vol. 21, No. 11, pp. 1020-1028, Nov. 2011. <http://dx.doi.org/10.5050/KSNVE.2011.21.11.1020>.
- [13] J. K. Kim, K. B. Lee, and S. G. Hong, "ECG-based Biometric Authentication Using Random Forest", Journal of the Institute of Electronics and Information Engineers, Vol. 54, No. 6, pp. 100-105, Jul. 2017.
- [14] S. J. Kim, J. H. Lee, D. H. Wang, and I. S. Lee, "Vehicle Simulator and SOC Estimation of Battery using Artificial Neural Networks", Journal of KIIT, Vol. 19, No. 5, pp. 51-62, May 2021. <http://dx.doi.org/10.14801/jkiit.2021.19.5.51>.
- [15] M. Sokolova, G. Lapalme, "A systematic analysis of performance measures for classification tasks", Information Processing and Management, Vol. 45, No. 4, pp. 427-437, Jul. 2009. <https://doi.org/10.1016/j.ipm.2009.03.002>.

### 저자소개

왕 동 훈 (Dong-Hun Wang)



2021년 2월 : 경북대학교  
전자공학부(공학사)  
2021년 3월 ~ 현재 : 경북대학교  
전자전기공학부(공학석사과정)  
관심분야 : 리튬 이온 배터리  
SOC추정 및 유도전동기  
고장진단

이 종 현 (Jong-Hyun Lee)



2018년 2월 : 경북대학교  
전자공학부(공학사)  
2020년 2월 : 경북대학교  
전자전기공학부(공학석사)  
2020년 3월 ~ 현재 : 경북대학교  
전자전기공학부(공학박사과정)  
관심분야 : 리튬 이온 배터리 SOC

및 SOH 추정

김 시 진 (Si-Jin Kim)



2020년 2월 : 영진전문대학교  
전자정보통신공학과(공학사)  
2020년 3월 ~ 현재 : 경북대학교  
전자전기공학부(공학석사과정)  
관심분야 : 리튬 이온 배터리  
SOC추정

김 민 찬 (Min-Chan Kim)



2021년 2월 : 경북대학교  
전자공학부(공학사)  
2021년 9월 ~ 현재 : 경북대학교  
전자전기공학부(공학석사과정)  
관심분야 : 리튬 이온 배터리  
SOC추정

이 인 수 (In-Soo Lee)



1986년 : 경북대학교  
전자공학부(공학사)  
1989년 : 경북대학교  
전자공학부(공학석사)  
1997년 : 경북대학교  
전자공학부(공학박사)  
1997년 3월 ~ 2008년 2월 :  
상주대학교 전자 및 전기 공학 교수  
2005년 8월 ~ 2007년 1월 : 미국 샌디에고 주립대학교  
방문학자  
2008년 3월 ~ 2014년 10월 : 경북대학교 전자전기공학부  
교수  
2014년 11월 ~ 현재 : 경북대학교 전자공학과 교수  
관심분야 : 배터리 SOC 및 SOH 추정, 시스템 고장진단  
및 고장허용제어, 신경회로망 기반 지능제어, 지능형  
센서 시스템