

## Research on Reactor Neutrino Event Selection by Using the Machine Learning Technique

Chang Dong SHIN · Kyung Kwang JOO\* · Dong Ho MOON<sup>†</sup>

Institute for Universe & Elementary Particles, Department of Physics,  
Chonnam National University, Gwangju 61186, Korea

Myoung Youl PAC<sup>‡</sup>

Institute of High Energy Physics, Dongshin University, Naju 58245, Korea

Junghwan GOH<sup>§</sup>

Department of Physics, Kyung Hee University, Seoul 02447, Korea

(Received 12 June 2019 : revised 01 July January 2019 : accepted 02 July 2019)

For the next-generation massive neutrino experiments, selecting a signal in the background events is very important. To do this, we investigated the results of applying a machine learning technique to the selection of neutrino signals. The neutrino signal after inverse beta decay and the background events in a gadolinium-loaded liquid scintillation detector were reproduced by using Monte Carlo simulations. The inverse beta decay process is well-known and has relatively high statistical quantities for this simulation. In this study, an efficiency of signal selection through machine learning was obtained, and in this paper several results are briefly described. Finally, the machine learning technique is expected to become an important tool for use in the next-generation neutrino experiment.

PACS numbers: 13.15.+g, 07.05.Tp, 29.40.Mc

Keywords: Machine learning, Neutrino, Liquid scintillator, Neutrino mixing angle, Next generation neutrino detector

## 기계학습을 이용한 원자로 중성미자 선별 연구

신창동 · 주경광\* · 문동호<sup>†</sup>

전남대학교 물리학과, 우주소립자 연구소, 광주광역시 61186, 대한민국

박명렬<sup>‡</sup>

동신대학교 고에너지물리 연구소, 나주시 58245, 대한민국

고정환<sup>§</sup>

경희대학교 물리학과, 서울특별시 02447, 대한민국

(2019년 6월 12일 받음, 2019년 7월 01일 수정본 받음, 2019년 7월 02일 개재 확정)

차세대 대형 중성미자 검출 실험에서 배경사건과 원하는 신호를 효율적으로 선별하는 것은 매우 중요하다. 이를 위해서 현재 유용하게 사용되고 있는 분석 기술의 하나인 기계학습을 사용하여 중성미자

신호 선별에 적용하였을 때의 결과를 살펴보고자 한다. 이를 위해 비교적 특징이 잘 알려지고, 상대적으로 통계량이 높은 원자로 중성미자의 역베타 붕괴 반응 이후 신호와 배경사건들을 몬테카를로 시뮬레이션을 통하여 재현하고, 기계학습을 통한 신호선별 효율을 확인하였다. 최종적으로는 향후 차세대 중성미자 실험에서 중요한 도구로 사용될 수 있을 것으로 기대한다.

PACS numbers: 13.15.+g, 07.05.Tp, 29.40.Mc

Keywords: 기계학습, 중성미자, 액체섬광검출용액, 진동변환상수, 차세대 중성미자 검출기

## I. 서 론

최근 입자물리 연구는 새로운 현상 발견을 위하여 높은 에너지 혹은 많은 수의 관측 사건을 만드는 방향으로 진행되고 있다. 이러한 변화에 맞추어 원하는 신호를 선별하고, 배경사건을 효율적으로 제거하기 위한 방법들 또한 새롭게 개발되고 있다. 그리고 물리뿐만이 아닌 다른 분야에서도 다양하고 많은 양의 데이터가 누적됨에 따라 기계학습이 폭넓게 활용되고 있다 [1]. 또한 기계학습은 현재 고에너지 입자 충돌 물리 분야에서도 유용하게 적용되고 있고, 다른 물리 분야에서도 서서히 데이터 분석에 활용할 것으로 예측된다.

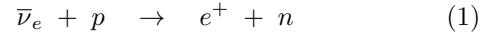
이와 더불어 표준모형 (Standard Model)에서 기본 입자 중의 하나인 중성미자는 최근 활발한 연구가 진행되어 왔고, 특히 진동변환상수는 상당히 정밀하게 측정하게 되었다. 하지만 직간접적인 관측 결과로부터 아직 잘 알려지지 않은 경입자 영역에서의 대칭성 깨짐, 중성미자 간의 질량 순서, 비활성 중성미자 존재여부 등을 주제로 한 표준 모형을 넘어서 새로운 여러 실험들이 계획 되고 있고, 혹은 건설이 진행 중에 있다 [2,3]. 이러한 새로운 사실들을 입증하기 위해서는 기존에 측정된 값들을 보다 정밀하게 측정하는 것이 필요하다. 또한 중성미자들의 기본적인 특징인 작은 산란 단면적으로 인한 적은 양의 신호를 수많은 배경사건들에서 선별하기 위한 기술이 매우 중요하다. 그러므로 본 논문에서는 기계학습을 통해 중성미자 신호와 배경사건 구별에 대한 연구를 진행하였다. 또한 비교적 그 특징이 잘 알려지고, 중성미자 중에서 상대적으로 많은 통계치를 가질 수 있는 원자로 중성미자 측정 실험을 전제로 진행하였다.

## II. 원자로 중성미자 및 기계학습

### 1. 원자로 중성미자와 역베타 붕괴반응

본 논문에서 시뮬레이션 한 중성미자는 원자로에서 방출되는 반전자 중성미자이고, 이를 검출하기 위해 액체섬광검

출용액이 채워진 검출기를 사용하였다. 반전자 중성미자가 검출기에 도달하면 검출기내의 양성자와 충돌하여 역베타 붕괴 반응을 일으킨다. 역베타 붕괴 시식 (1)과 같이 양전자와 중성자가 방출된다.



양전자 ( $e^+$ )는 거의 즉시 쌍소멸 하여 1–10 MeV의 빛을 방출하고 (S1, prompt), 중성자 ( $n$ )의 경우는 액체섬광 검출용액에 첨가된 물질에 따라, 그 반응이 달라진다 [4, 5]. 현재 원자로 중성미자 검출을 위해 건설된 일반적인 검출기는 원통형의 모양으로 필요에 따라 몇 개의 층으로 구성된다. 원통의 가장 안쪽은 약 0.1%의 가돌리늄 (Gd)이 용해된 액체섬광검출용액을 중성미자 표적으로 사용하고, 그 외부는 가돌리늄이 용해되지 않은 액체섬광검출용액이 채워진다 [6,7]. 이 경우 중성미자 표적에서는 약 85%가 평균적으로 30  $\mu$ s 이후에 가돌리늄에 포획되어 약 8 MeV의 빛을 방출한다. 이를 제외한 중성미자 표적 바깥 영역까지의 대부분의 이벤트는 역베타 붕괴 이후 중성자가 평균적으로 200  $\mu$ s 이후에 수소에 포획되어 2.2 MeV의 빛을 방출한다 (S2, delayed). 일반적으로 가돌리늄에 포획된 신호의 경우 중성자 신호의 높은 에너지 및 양전자 신호와의 짧은 시간 차이의 특징으로 인해 배경사건을 보다 쉽게 제거할 수 있다. 하지만 수소에 포획된 경우 중성자 신호가 2.2 MeV로 검출기 주변의 바위로부터 생성되어 들어오는 배경사건 신호와 에너지 영역이 비슷하여 신호 선별에 상당한 어려움이 있다. 그러므로 보다 신호 선별이 어려운 수소 포획 신호에 대하여 연구를 진행하였다.

### 2. 기계학습 방법 및 도구

본 논문의 연구를 위해서 사용한 분석도구는 CERN에서 개발되어 입자물리분야에서 유용하게 사용되고 있는 ROOT에 포함되어 있다. Toolkit for Multivariate Data Analysis (TMVA)라고 불리는 이 도구는 다양한 다변수 분류 알고리즘을 제공하는데 이 논문에서는 그 중에서 Multilayer Perceptron (MLP) 방법을 사용하였고 또한 전통적

\*E-mail: kkjoo@chonnam.ac.kr

†E-mail: dhmoon@jnu.ac.kr

‡E-mail: pac@dsu.ac.kr

§E-mail: jhgo@khu.ac.kr

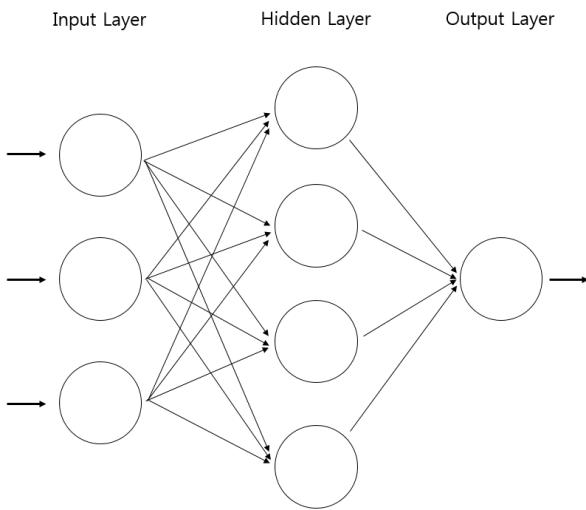


Fig. 1. The scheme of multi-layer perceptron. This method is based on the artificial neural network. Generally, it consists of the three layers such as input, hidden and output.

인 Boosted Decision Tree (BDT) 방법으로 교차확인도 하였다. MLP는 인간 두뇌의 신경망에서 착안된 것으로, 신경세포인 뉴런들이 연결된 형태로 구성된다. Figure 1은 일반적인 인공신경망의 구조를 도식화한 것이다. 인공신경망은 입력층 (Input layer), 은닉층 (Hidden layer), 출력층 (Output layer)으로 구성된다. 기본적으로 학습에 사용되는 데이터가 입력층을 통해서 들어가면, 은닉층에서 다양한 연산과정을 거치게 된 후, 입력된 데이터들을 출력층에서 어떠한 한 개의 값으로 나타낼 수 있다. 한 개의 은닉층에는 여러 개의 노드들이 존재할 수 있고, 각각의 노드 사이의 연결에서는 다른 값의 가중치들을 적용할 수 있다. 또한 Fig. 1에서는 한 개의 은닉층만 존재하지만, 이 은닉층은 여러 개가 존재할 수 있고, 이에 따라 기계학습을 위한 연산시간 또한 길어지게 된다.

### III. 기계학습 결과

#### 1. 학습 데이터

학습에 사용된 데이터는 중성미자 신호와 배경사건으로 구분된다. 중성미자 신호는 앞에서 설명된 것과 같이 반전자 중성미자가 검출기에 도달하여 역베타 붕괴반응 이후 생성된 양전자와 중성자 신호를 재현하였다. 중성자가 수소에 포획된 중성미자 신호를 측정 할 때, 가장 큰 비중을 차지하는 배경사건은 검출기 주변을 둘러싸는 바위에서부터 나오는 자연방사능이다. 일반적으로 중성미자 검출기는

우주선에 의한 배경사건을 줄이기 위하여 지하에 건설되고, 이에 따라 주변 바위에 포함된  $^{40}\text{K}$ ,  $^{232}\text{Th}$ ,  $^{238}\text{U}$  등의 방사선 동위원소에서 방출된 빛이 검출기에 들어온다 [8]. 이 빛들 중 허용된 짧은 시간 내에 우연히 두 개가 들어올 경우 중성미자로부터 생성된 양전자와 중성자 쌍을 흉내낸다. 중성미자 신호의 경우는 S1과 S2 사이에는 중성자의 포획시간에 따른 연관관계가 존재한다. 하지만 배경사건의 경우에는 우연한 두 빛 사이의 결합으로, S1과 S2 사이에 어떠한 연관관계도 없다.

중성미자 신호의 학습데이터를 재현하기 위해서 역베타 붕괴 후 생성되는 양전자와 중성자의 검출기내에서 상호 작용에 대한 시뮬레이션이 필요하다. 입력되는 양전자와 중성자의 운동량 및 위치정보는 원자로 중성미자의 방출량과 역베타 붕괴 반응의 산란단면적을 이용하여 계산된다 [9]. 그 이후 양자의 쌍소멸, 중성자 포획이 발생하는 위치 및 시간은 GEANT4를 기반으로 한 GLG4SIM을 이용하여 기본적인 원통형 검출기에 대하여 중성미자 신호를 시뮬레이션 하였다 [10].

Figure 2는 기계학습에 사용된 변수들의 분포를 보여준다. 두 재현된 데이터의 변수는 S1과 S2 사이의 시간차이 ( $\Delta T$ ) 및 거리 ( $\Delta R$ ), 각각의 에너지이다 ( $E_{S1}, E_{S2}$ ). 중성미자 신호의 에너지의 경우 앞에서 설명한 분포를 보였고, 배경사건의 경우 낮은 에너지부터 약 3 MeV 까지 감소하는 분포를 보인다. 또한 두 신호의 시간차이와 거리차이의 경우 역시 서로 연관관계에 의해 짧은 시간 및 거리를 보여준다. 그에 비해 배경사건의 경우는 무작위로 쌍을 이루기 때문에 균일한 시간분포와 검출기의 크기에 의해서만 제한되는 거리 분포를 가진다. 그 외에 같이 그려진 중성자가 가돌리늄에 포획된 경우는 대부분의 특성이 수소에 포획된 경우와 비슷한 범위에 분포하므로 배경사건처럼 기여할 수 있다. 하지만 Fig. 2에서 보듯이 중성자 신호의 에너지가  $\sim 8$  MeV로 명확하게 구분되어서 쉽게 구분할 수 있다.

#### 2. 신호 선별 및 배경사건 제거 효율

기계학습은 MLP방법을 사용하여 5만개의 중성미자 신호 및 배경사건을 학습하여 진행하였다. 기본적으로 배경사건 제거를 위해서는 S1과 S2의 연관성을 나타내는 시간차이와 위치 차이가 가장 효율적으로 작동할 것으로 기대한다. 이에 따라 두 개의 변수에 대한 학습결과를 확인하고, 추가적으로 S1과 S2의 특징을 나타내는 에너지 변수를 같이 학습하였을 시의 결과를 비교하였다. 또한 MLP 방법에 사용된 은닉층의 수는 1로 동일하고, 노드의 수는 에너지 변수를 사용하지 않은 경우 7, 에너지 변수를 사용한 경우

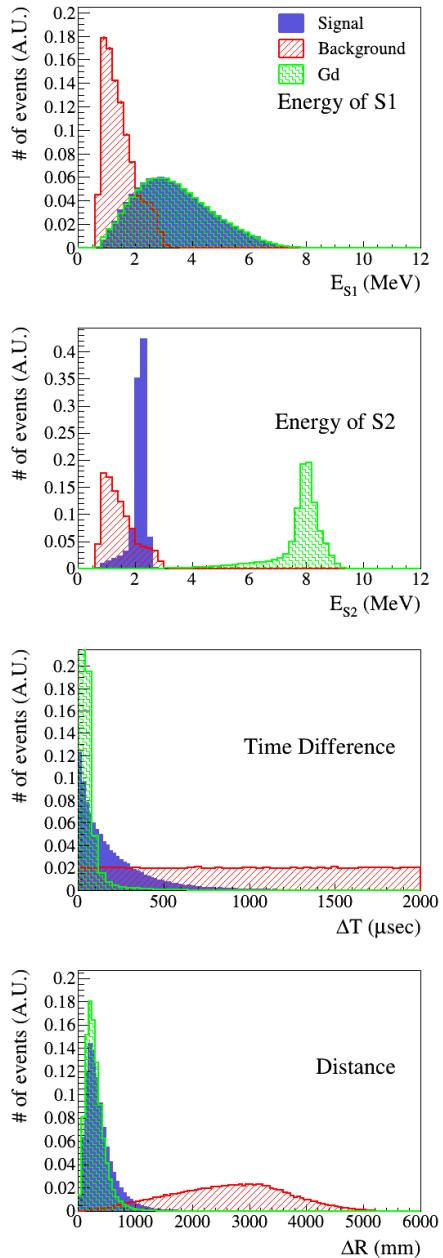


Fig. 2. (Color online) Input variables for machine learning. S1 and S2 represent the energy distribution of prompt and delayed signals,  $\Delta T$  is time difference between S1 and S2, and  $\Delta R$  indicate the distance between S1 and S2.

9이다. Figure 3은 중성미자의 검출효율에 대한 배경사건의 제거 효율을 나타낸다. 이 결과는 배경사건의 시간차이 분포에 의존하고, 현재 결과는 중성미자 신호의 시간차이 분포에 근거하여 2000  $\mu$ s 까지를 기준으로 하였다. 에너지 분포를 사용한 경우 및 사용하지 않은 경우를 비교하였을 때, 명확히 많은 변수를 사용한 경우가 배경사건을 좀 더 효율적으로 제거하는 것을 보여주고 있다. 하지만 시간 차

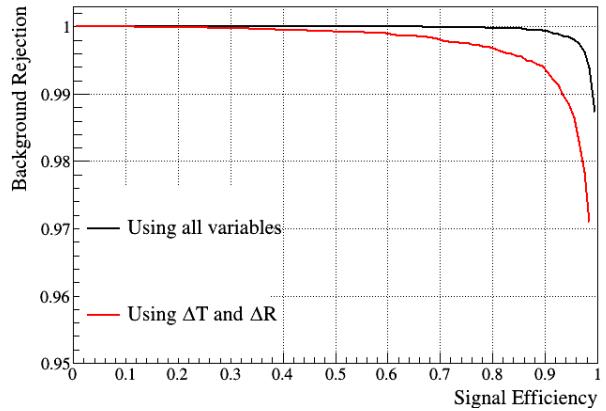


Fig. 3. (Color online) Background rejection and signal efficiency by machine learning. The black line is the result using all variables and the red line is the result using only  $\Delta T$  and  $\Delta R$  without energy variables. Both results show that background can be removed efficiently by machine learning.

이와 위치 차이만 사용한 경우에도 배경사건의 특징이 중성미자에 비하여 비교적 명확하여, 배경사건이 효과적으로 제거되는 것을 확인할 수 있다.

Figure 4는 두 가지 학습결과를 적용한 중성미자 신호의 양전자의 에너지 분포를 보여주고, 각 성분의 분포를 비교하기 위해 전체 양은 조절되었다. 이 경우 각 신호의 에너지를 사용하여 학습한 경우 중성미자 신호에서 제거된 이벤트가  $\sim 3$  MeV 이하의 낮은 영역에 분포하는 것을 볼 수 있다. 이것은 배경사건의 에너지 특성이 반영된 결과라고 할 수 있다. 이 경우 실제 중성미자 신호의 경우도 특정 에너지의 이벤트들이 주로 제거되기 때문에 최종적으로 관측되는 에너지 분포가 변경되어, 검출된 중성미자의 에너지를 이용한 분석을 진행할 경우 결과에 영향을 줄 수 있다. 특히 중성미자의 진동변환 상수 중에서 중성미자간의 질량차이를 측정하는 경우에는 양전자의 에너지 분포를 사용하므로 영향을 준다. 이는 분석에 사용되는 변수의 경우 학습 데이터로 사용하는 것을 지양해야 하는 것을 의미한다. 그러므로 본 논문에서 확인한 두 가지 방법 중 거리차이와 시간차이만 사용한 결과가 배경사건 제거 측면에서는 효율이 낮더라도, 적합한 기계학습 방법이 될 수 있다.

#### IV. 요 약

원자로로부터 생성된 반전자 중성미자 신호와 배경사건을 사용한 기계학습 결과가 효과적으로 작동함을 확인하였다. 물론 현재 사용된 중성미자와 배경사건의 경우 특징이 명확하여 기계 학습을 적용하지 않더라도 많은 배경사건을

## 감사의 글

이 논문은 전남대학교 학술연구비 (과제번호 : 2018-3471), 미래창조과학부의 재원으로 한국연구재단의 기초 연구사업 (NRF-2016RIA2B4008925, 2019R1A2B5B01070451, 2019R1I1A1A01059548, 2016R1D1A3B02010606)과 Samsung Science & Technology Foundation (SSTF-BA1402-06), 2018년도 경희대학교 연구비 지원 (KHU-20180939)의 지원을 받아 수행된 연구입니다.

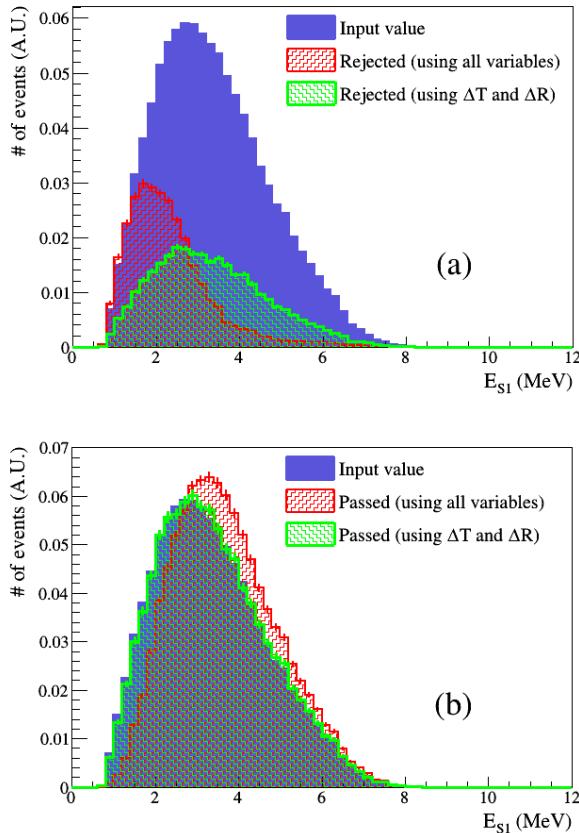


Fig. 4. (Color online) Background rejection (a) and signal efficiency (b) by machine learning. The red line is the result using all variables and the green line is the result using only  $\Delta T$  and  $\Delta R$  without energy variables. Both results show that background can be removed efficiently by machine learning.

제거할 수 있다. 하지만 실제 실험환경에서 신호와 배경사건의 비율을 고려한다면, 최종적으로 선별된 데이터의 중성미자 신호와 배경사건의 비율을 향상시킬 수 있을 것이다. 현재 연구에서는 기계학습을 중성미자 검출에서도 충분히 적용가능 하다는 것에 의의를 두며, 향후 신호의 특징이 명확하지 않은 대규모 차세대 중성미자 실험에서 중요한 도구로 사용될 수 있을 것으로 기대한다.

## REFERENCES

- [1] A. Radovic, M. Williams, R. David and M. Kagan *et al.*, *Nature* **560**, 41 (2018).
- [2] S. Fukasawa, M. Ghosh, and O. Yasuda, *Phys. Rev. D* **95**, 055005 (2017).
- [3] F. An, G. An, Q. An and V. Antonelli *et al.*, *J.J. Phys. G: Nucl. Part. Phys.* **43**, 3 (2016).
- [4] B. R. Kim, S. C. Kim, K. K. Joo and C. D. Shin *et al.*, *J. Korean Phys. Soc.* **66**, 5, 768 (2015).
- [5] Y. S Park, Y. M. Jang and K. K. Joo, *Rev. Sci. Instrum.* **89**, 043302 (2018).
- [6] J. K. Ahn, S. Chebotaryov, J. H Choi and S. Choi, *et al.*, (RENO Collaboration), *Phys. Rev. Lett.* **108**, 191802 (2012).
- [7] J. S Park, J. Lee, I. S. Yeo and W. Q. Choi *et al.*, (RENO Collaboration), *Nucl. Instrum. Meth.* **707**, 45 (2013).
- [8] M. M. Makweba, E. Holm, *Sci. Total Environ.* **133**, 1, p99 (1993).
- [9] P. Vogel and J. Beacom, *Phys. Rev. D* **60**, 053003 (1999).
- [10] CERN Program Library Long Writeup, No.W5013 (1993).