

저 사양 컴퓨팅 환경에서 기계학습 성능 향상을 위한 주성분 분석 기반 데이터 전처리 정책

이현섭*

백석대학교 컴퓨터공학부 교수

A Principal Component Analysis Based Data Preprocessing Policy for Improving Machine Learning Performance in Low Specification Computing Environments

Hyun-Seob Lee*

Professor, Division of Computer Engineering, Baekseok University

요약 최근 다양한 분야에서 인공지능 적용을 통한 자동화 환경 구축을 위한 연구가 활발히 진행되고 있다. 그러나 기계학습을 적용하기 위한 고 사양 컴퓨팅 환경 구축에는 상당한 비용이 소요된다. 이에 따라 저 사양 환경에서 인공지능을 효율적으로 적용하기 위한 연구의 필요성이 대두되고 있다. 저 사양 컴퓨팅 환경에서 고차원 데이터에 대한 기계학습은 메모리 제약과 연산 비용으로 인해 심각한 성능 저하를 겪는다. 본 연구는 이러한 문제를 해결하기 위해 주성분 분석(Principal Component Analysis, PCA)을 활용한 차원 축소 기법을 제안한다. 제안하는 방법은 고차원 데이터에 대한 기계학습 수행 전 PCA를 통한 전처리 과정을 추가하고, 이를 통해 정제된 데이터를 활용하여 기계학습을 수행한다. 이로써 제한된 컴퓨팅 자원 환경에서도 기계학습 모델의 성능과 효율성을 동시에 개선할 수 있다. 실험을 통해 동일한 환경에서 PCA 기반 전처리가 다차원 데이터의 차원을 효과적으로 축소하면서도 데이터의 주요 분산 정보를 보존함을 확인하였다. 그 결과, 평균 학습 시간이 96.6% 감소하였고 메모리 사용량은 99.7% 절감되었으며, 모델의 예측 정확도는 유사한 수준으로 유지되는 것을 확인하였다. 본 연구는 임베디드 시스템이나 엣지 컴퓨팅과 같이 제한된 하드웨어 자원을 가진 환경에서도 효율적인 기계학습 구현이 가능함을 입증하였다. 이는 고성능 컴퓨팅 인프라에 대한 접근이 제한된 연구자나 개발자에게 실용적인 솔루션을 제공할 수 있을 것으로 기대된다.

주제어 : 주성분 분석, 기계학습, 차원 축소, 저 사양 컴퓨팅, 데이터 전처리

Abstract Research is actively underway to establish automated environments through the application of artificial intelligence across various fields. However, building high-specification computing environments for applying machine learning incurs significant costs. Consequently, the need for research on efficiently applying AI in low-specification environments has emerged. Machine learning on high-dimensional data in low-specification computing environments suffers severe performance degradation due to memory constraints and computational costs. This study proposes a dimension reduction technique utilizing Principal Component Analysis (PCA) to address these issues. The proposed method adds a preprocessing step using PCA before performing machine learning on high-dimensional data, enabling machine learning to be performed using the refined data. This simultaneously improves the performance and efficiency of machine learning models even in environments with limited computing resources. Experiments confirmed that PCA-based preprocessing effectively reduces the dimensions of multidimensional data while preserving key variance information under identical conditions. Consequently, average training time decreased by 96.6%, memory usage was reduced by 99.7%, and model prediction accuracy remained at a comparable level. This research demonstrates that efficient machine learning implementation is possible even in environments with limited hardware resources, such as embedded systems or edge computing. This is expected to provide a practical solution for researchers and developers with limited access to high-performance computing infrastructure.

Key Words : Principal Component Analysis, Machine Learning, Dimensionality Reduction, Low Specification Computing, Data Preprocessing

This paper was supported by 2026 Baekseok University Research Fund

*교신저자 : 이현섭(hyunseob@bu.ac.kr)

접수일 2026년 01월 31일 수정일 2026년 05월 03일 심사완료일 2026년 06월 05일

1. 서론

최근 인공지능 기술의 급속한 발전과 함께 다양한 산업 분야에서 기계학습을 활용한 자동화 시스템 구축이 확대되고 있다. 제조업의 품질 검사, 의료 영상 분석, 금융 거래 모니터링, 농업 생산성 관리 등 거의 모든 영역에서 데이터 기반 의사결정과 예측 모델의 중요성이 강조되고 있다. 이러한 추세에 따라 기계학습 모델의 성능을 최대화하기 위해 대용량의 고차원 데이터를 처리하는 경우가 증가하고 있다.

그러나 고성능 기계학습 모델의 학습과 운영을 위해서는 고사양의 컴퓨팅 자원이 필수적이다. 특히 GPU를 포함한 고성능 하드웨어, 대용량 메모리, 그리고 이를 지원하는 인프라 구축에는 상당한 초기 투자 비용과 유지보수 비용이 소요된다. 이는 중소기업이나 연구 기관, 개발도상국의 연구자들에게 기술 접근성의 장벽으로 작용하고 있다. 또한 클라우드 컴퓨팅 서비스를 이용하는 경우에도 지속적인 사용 비용이 발생하여 경제적 부담이 될 수 있다. 더욱이 임베디드 시스템, 사물인터넷(IoT) 기기, 모바일 기기, 엣지 컴퓨팅 환경과 같이 제한된 컴퓨팅 자원을 가진 환경에서의 기계학습 적용 수요가 증가하고 있다. 이러한 환경에서는 실시간 처리와 저전력 운영이 요구되지만, 전통적인 기계학습 접근법은 메모리 제약과 연산 능력의 한계로 인해 심각한 성능 저하를 겪게 된다. 특히 고차원 데이터를 다루는 경우 메모리 부족, 과도한 학습 시간, 그리고 과적합(overfitting) 문제가 발생할 가능성이 높아진다. 이러한 문제를 해결하기 위해 저사양 환경에서도 효율적으로 기계학습을 수행할 수 있는 방법론에 대한 연구가 필요하다. 본 연구는 차원 축소 기법 중 하나인 주성분 분석(Principal Component Analysis, PCA)[1-2]을 활용하여 저사양 컴퓨팅 환경에서 기계학습 모델의 성능과 효율성을 동시에 향상시킬 수 있는 방법을 제시한다.

본 연구의 목적은 저사양 컴퓨팅 환경에서 PCA 기반 차원 축소를 통해 기계학습의 효율성을 개선하고, 동시에 모델의 예측 정확도를 유지할 수 있음을 실증적으로 검증하는 것이다. 구체적인 연구 목표는 다음과 같다.

첫째, 고차원 데이터에 대한 PCA 적용이 저사양 환경에서 학습 시간과 메모리 사용량에 미치는 영향을 정량적으로 분석한다. 둘째, 차원 축소 후 기계학습 모델의 예측 정확도 변화를 측정하여 PCA의 효과성을 평가한다. 셋째, 다양한 데이터셋과 기계학습 알고리즘에 대해 제안하는 방법의 적용 가능성을 검증한다. 본 연구의 범

위는 감독학습(supervised learning) 기반의 분류 및 회귀 문제를 대상으로 하며, 다양한 차원의 데이터셋에 대해 PCA 전처리의 효과를 비교 분석한다. 실험은 제한된 메모리와 연산 능력을 가진 저사양 하드웨어 환경에서 수행되며, 원본 데이터를 사용한 경우와 PCA를 적용한 경우의 성능을 비교한다.

본 논문의 구성은 다음과 같다. 2장에서는 차원 축소 기법과 주성분 분석에 대한 이론적 배경 및 관련 연구를 살펴본다. 3장에서는 제안하는 PCA 기반 전처리 방법론과 실험 설계를 상세히 기술한다. 4장에서는 다양한 데이터셋에 대한 실험 결과를 제시하고 분석한다. 마지막으로 5장에서는 연구의 결론과 향후 연구 방향을 제시한다.

2. 배경

2.1 고차원 데이터의 문제점

현대의 기계학습 문제에서 다루는 데이터는 점점 더 고차원화되고 있다. 이미지 데이터, 유전자 발현 데이터, 텍스트 데이터, 센서 데이터 등은 수백에서 수만 개 이상의 특징(feature)을 포함할 수 있다. 고차원 데이터는 다음과 같은 문제를 야기한다. 첫째, 차원의 저주(curse of dimensionality)[3] 현상이 발생한다. 특징의 수가 증가함에 따라 데이터 공간의 부피가 기하급수적으로 증가하여, 동일한 밀도를 유지하기 위해 필요한 샘플 수가 급격히 증가한다. 이는 모델의 일반화 성능을 저하시키고 과적합을 유발할 수 있다. 둘째, 계산 복잡도가 크게 증가한다. 대부분의 기계학습 알고리즘은 특징 수에 따라 계산 시간이 증가하며, 특히 거리 기반 알고리즘이나 커널 방법의 경우 이러한 영향이 더욱 두드러진다. 저사양 환경에서는 이러한 계산 부담이 실용적인 적용을 어렵게 만든다. 셋째, 메모리 요구량이 증가한다. 고차원 데이터를 메모리에 적재하고 처리하는 과정에서 상당한 메모리 공간이 필요하며, 제한된 메모리를 가진 환경에서는 데이터 전체를 처리하지 못하는 문제가 발생할 수 있다.

2.2 차원 축소 기법

차원 축소(dimensionality reduction)는 고차원 데이터의 문제를 해결하기 위한 핵심 전처리 기법이다. 차원 축소는 원본 데이터의 중요한 정보를 최대한 보존하

면서 특징의 수를 줄이는 것을 목표로 한다. 차원 축소 기법은 크게 특징 선택(feature selection)과 특징 추출(feature extraction)로 구분된다. 특징 선택은 원본 특징 중에서 가장 중요한 부분 집합을 선택하는 방법이다. 필터 방법(filter method), 래퍼 방법(wrapper method), 임베디드 방법(embedded method) 등이 있으며, 각각 통계적 기준, 모델 성능, 학습 과정과의 결합 등 다양한 기준을 활용한다. 특징 추출은 원본 특징들의 조합을 통해 새로운 특징 공간을 생성하는 방법이다. 주성분 분석(PCA), 선형 판별 분석(Linear Discriminant Analysis, LDA), 독립 성분 분석(Independent Component Analysis, ICA), t-SNE, UMAP 등이 대표적인 기법이다. 이 중 PCA는 가장 널리 사용되는 선형 차원 축소 기법으로, 데이터의 분산을 최대화하는 방향으로 새로운 축을 찾아 변환한다.

2.3 주성분 분석(PCA)

주성분 분석(PCA)은 데이터 분석을 위해 연구된 통계적 기법으로 상관관계가 있는 고차원 데이터를 상관관계가 없는 저차원 데이터로 변환하는 선형 변환 기법이다. PCA의 핵심 아이디어는 데이터의 분산이 최대가 되는 방향을 찾아 새로운 좌표축으로 설정하는 것이다. 첫 번째 주성분은 데이터 분산이 가장 큰 방향이며, 두 번째 주성분은 첫 번째 주성분과 직교하면서 남은 분산이 가장 큰 방향이 된다. 이러한 과정을 반복하여 데이터의 차원만큼 주성분을 얻을 수 있다. 수학적으로 PCA는 데이터의 공분산 행렬에 대한 고유값 분해(eigenvalue decomposition)[4] 또는 특이값 분해(singular value decomposition, SVD)[5]를 통해 수행된다. 고유값이 큰 순서대로 대응하는 고유벡터들이 주성분이 되며, 상위 k 개의 주성분만을 선택함으로써 차원 축소가 이루어진다. PCA의 장점은 다음과 같다. 첫째, 계산이 비교적 단순하고 효율적이다. 둘째, 선형 변환이므로 해석이 용이하다. 셋째, 데이터의 주요 분산 정보를 보존한다. 넷째, 특징 간의 상관관계를 제거하여 중복 정보를 감소시킨다.

2.4 저사양 환경에서의 기계학습 관련 연구

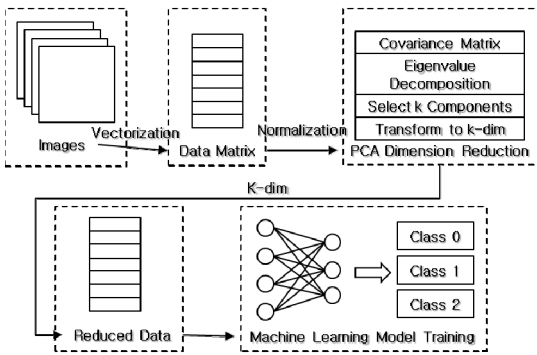
저사양 컴퓨팅 환경에서의 효율적인 기계학습에 대한 연구는 여러 방향에서 진행되어 왔다. 모델 압축(model compression)[6], 경량화 모델 설계, 양자화(quantization)[7], 지식 증류(knowledge distillation)[8] 등의 기법들이

제안되었다. 그러나 이러한 연구들은 주로 모델 자체의 복잡도를 줄이는 데 초점을 맞추고 있다. 데이터 전처리 관점에서의 접근은 상대적으로 덜 주목받았으나, 최근 엣지 컴퓨팅과 IoT 환경의 확대와 함께 중요성이 증가하고 있다. 일부 연구에서는 차원 축소를 통한 효율성 개선을 다루었으나, 저사양 환경에 특화된 체계적인 분석은 부족한 실정이다. PCA와 기계학습의 결합에 관한 기존 연구들은 주로 고사양 환경에서의 정확도 향상이나 시각화에 초점을 맞추었다. 반면 본 연구는 저사양 환경에서의 실용성과 효율성에 중점을 두어, 제한된 자원 환경에서 PCA가 학습 시간, 메모리 사용량, 그리고 모델 정확도에 미치는 영향을 종합적으로 분석한다는 점에서 차별성을 갖는다.

3. 주성분 분석 기반 데이터 전처리 정책

3.1 핵심 아이디어

본 논문에서 제안하는 방법의 핵심 아이디어는 저사양 컴퓨팅 환경에서 기계학습의 효율성과 실용성을 동시에 확보하기 위해 주성분 분석(PCA)을 전처리 단계로 활용하는 것이다. 기존의 기계학습 파이프라인에 PCA 기반 차원 축소 과정을 삽입함으로써, 고차원 데이터가 가진 계산 복잡도와 메모리 부담을 근본적으로 감소시키면서도 데이터의 핵심 정보는 보존한다. 제안하는 접근법의 핵심은 다음 세 가지 원칙에 기반한다. 첫째, 선제적 복잡도 감소이다. 기계학습 모델에 데이터를 투입하기 전에 PCA를 통해 차원을 축소함으로써, 학습 과정에서 발생하는 계산 부담을 사전에 줄인다. 이는 사후적으로 모델을 경량화하는 접근법과 달리, 데이터 자체의 복잡도를 근원적으로 해결한다는 점에서 차별화된다. 둘째, 정보 보존적 압축이다. PCA는 무작위로 특징을 제거하는 것이 아니라, 데이터의 분산을 기준으로 중요도를 평가하여 핵심 정보를 보존한다. 상위 주성분들은 데이터의 대부분의 분산을 설명하므로, 적절한 수의 주성분을 선택하면 원본 데이터의 본질적 특성을 유지하면서도 차원을 크게 줄일 수 있다. 셋째, 자원 제약 환경 최적화이다. 저사양 환경의 제약 조건인 제한된 메모리와 연산 능력을 명시적으로 고려하여, PCA의 계산 비용을 포함하더라도 전체적인 학습 과정에서 순이익이 발생하도록 설계한다.



[Fig. 1] Key Idea

Fig. 1은 주성분 분석 기반 데이터 전처리 정책의 핵심 아이디어를 보여주고 있다. 그림과 같이 5단계로 구분된다. 1단계는 원본 데이터를 입력받는 단계다. 이 단계에서는 로딩한 데이터를 처리하기 위한 준비과정을 수행한다. 그다음 2단계는 데이터 행렬을 구성하는 단계다. 이 과정에서는 1차원 벡터로 변환하고 일정 범위로 스케일링하는 정규화 과정을 거친다. 3단계는 주성분 분석을 이용한 차원 축소과정이다. 이 단계에서는 각 특징의 평균을 계산하여 중심을 0으로 이동시킨다. 그다음 공분산 행렬을 계산하고 고유값이 큰 순서대로 정렬한다. 마지막으로 전체 분산의 90%이상에 해당하는 주성분을 k개 선택하여 차원을 축소 변환한다. 4단계는 선별된 k차원의 데이터를 선형 결합한다. 이 과정에서 데이터의 양을 감소시킨다. 마지막으로 5단계는 축소된 데이터 행렬을 이용하여 설계된 머신 러닝 모델에서 학습과정을 수행한다. 그림의 예에서는 선형 모델을 사용하였으나, 학습성능과 효율을 고려하여 적당한 모델을 선택하여 적용한다.

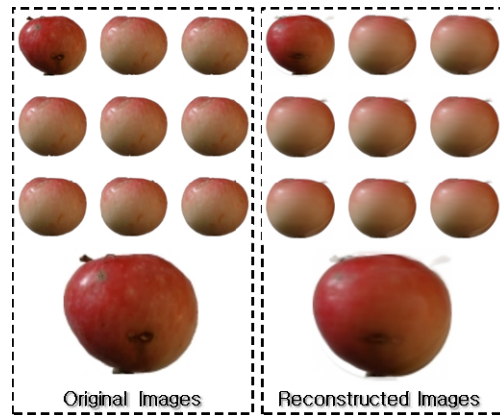
3.2 PCA 기반 차원 축소 정책

이 정책은 1차로 수집되어 벡터화 및 정규화 과정을 통해 전처리 된 데이터에 대해 PCA를 적용하여 차원을 축소한다. 이 과정은 다음의 하위 단계로 구성된다.

- 1) 데이터 중심화(Centering): 각 특징의 평균을 0으로 만들어 데이터를 중심화한다.
- 2) 공분산 행렬 계산: 중심화된 데이터의 공분산 행렬 $C = (1/n)X^T X$ 를 계산한다.
- 3) 고유값 분해: 공분산 행렬에 대한 고유값 분해를 수행하여 고유값 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$ 와 대응하는 고유벡터 v_1, v_2, \dots, v_d 를 구한다.
- 4) 주성분 선택: 누적 분산 기여율(cumulative

explained variance ratio)을 기준으로 k개의 주성분을 선택한다. 일반적으로 누적 분산 기여율이 85~95%에 도달하는 지점을 선택한다.

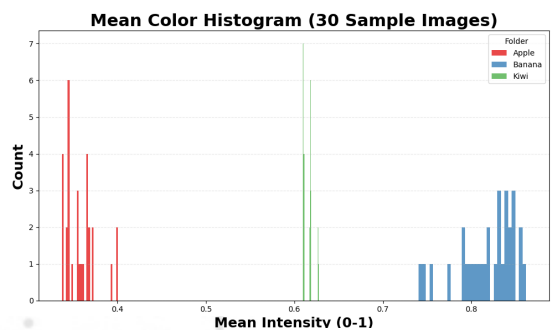
- 5) 차원 축소 변환: 선택된 k개의 고유벡터로 구성된 변환 행렬 $W = [v_1, v_2, \dots, v_k]$ 를 이용하여 원본 데이터를 $Z = XW \in R^{(n \times k)}$ 로 변환한다.



[Fig. 2] PCA Data Comparison

Fig. 2는 PCA로 선별된 데이터를 원본 이미지와 비교한 시각적 이미지 결과를 보여주고 있다. 그림에서는 9개의 샘플 이미지와 첫 번째 이미지의 확대된 이미지를 보여주고 있다. 왼쪽 그룹은 원본 이미지이고 오른쪽 그룹은 PCA 분석이 완료된 데이터를 이미지로 재구성하여 보여준 결과다. 시각화된 이미지 결과와 같이 데이터를 시각화한 결과는 오리지널 이미지와 유사한 결과를 얻을 수 있다. 그러나 질감과 일부 색상 등 인식을 하는데 큰 기여를 하지 않은 특징은 PCA를 통해 제거된 것을 확인할 수 있다.

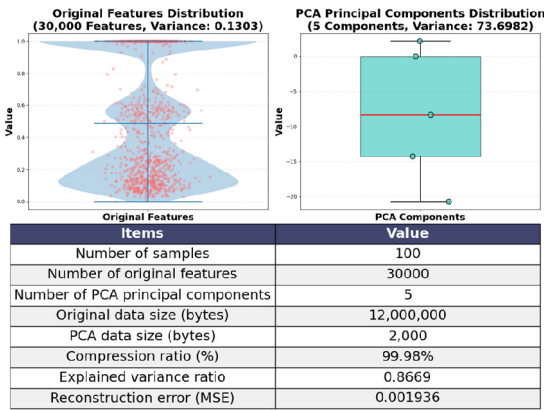
3.3 데이터 주성분 효과 분석



[Fig. 3] Color Histogram of 30 Sample Images

이미지 데이터는 다양한 특징을 가지고 있다. 그 중, 색상 값의 주요 성분을 분석하는 것만으로 동일 그룹 안에 이미지 분류가 가능하다.

Fig. 3은 Apple, Banana, Kiwi 이미지에 대해 10장씩 30개의 데이터에 대한 컬러 히스토그램에 대한 평균 분석한 결과다. 그림의 결과와 같이 단순히 색상 특성으로 이미지를 충분히 분류할 수 있다. 따라서 PCA를 통한 상위 주성분으로 이미지 분류가 가능하다.



[Fig. 4] PCA Data Distribution

Fig. 4는 100개의 이미지 샘플에 대한 5개의 주성분 분포를 보여주고 있는 결과다. 그림의 예제에서는 원본 이미지의 특징은 픽셀값 기준으로 30,000개였고, 그중 주성분은 빈도수 기준 5개를 선택하였다. 5개의 특징으로 압축했을 때 약 12MB의 크기를 2KB 이내로 줄일 수 있었다. 그러나 이러한 분석은 극단적인 효과를 판단하기 위하여 주성분을 극단적으로 줄였고, 데이터의 특성이 명백하게 분류된 데이터를 구분하기 위해서는 전체 데이터에 대한 분산 비율이 0.9 이상의 값을 갖도록 설정해야한다.

4. 실험 및 분석

본 절에서는 데이터에 대해 PCA 주성분 분석을 전처리하여 학습 하는 것이 미치는 영향을 확인하기 위해 간단한 실험을 통해 성능과 자원의 소모를 비교하였다. 실험에서 사용한 데이터는 Kaggle에 기계학습 연구자들을 위해 공개된 데이터 셋[9]을 사용하였다. 이 데이터는 237 종의 과일, 채소, 견과류, 씨앗을 촬영한 총 166,293 장의 이미지 데이터다. 이 중 124,716장은 기계학습 데이

터로 사용하였고, 41,577장은 데이터 분류 성공 여부를 확인하기 위한 검증용 데이터로 사용하였다. 실험을 한 환경은 AMD Ryzen Threadripper PRO 5975WX 32-Cores, 3600Mhz, 512GB RAM, RTX4090x4를 탑재한 윈도우 OS환경의 워크스테이션에서 진행하였다. 기계학습 모델은 Scikit-learn[10-13]에서 제공하는 기본 Logistic Regression 모델[14, 15]을 사용하였다.

Items	Original Data	PCA Data
Accuracy	94.29%	89.98%
Duration	1:02:56.80	0:02:08.05
Original Data of Training	14,272.61MB	
PCA Data of Training	47.58MB	
Original Data of Test	4,758.11MB	
PCA Data of Test	15.86MB	

[Fig. 5] Evaluation Results

Fig. 5는 기계학습을 수행한 평가 결과를 보여주고 있다. 그림의 결과와 같이 원본 데이터와 PCA 데이터를 이용한 기계학습 이후 정확도 테스트 결과는 94.29%와 89.98%로 원본 데이터를 이용하여 학습하는 것이 4.31% 높았다. 반면 기계학습에 소요된 시간은 각각 1:02:56.80초와 2:08.05초로 PCA데이터를 이용하여 학습을 하는 것이 원본 데이터를 사용하는 것보다 약 96.6% 향상되었다. 이 것은 PCA를 통해 차원을 축소하는 것은 학습 성능과 정확도 사이 트레이드오프(trade-off)임을 의미한다. 또한 자원 사용면에서도 트레이닝 데이터는 원본 데이터인 14.27GB가 47.58MB로 감소하였고, 테스트 데이터는 4.56GB가 15.86MB로 감소하여 약 99.7% 메모리 자원을 절약하는 효과가 있었다. 따라서 주성분 분석을 이용하여 차원을 축소하는 전처리 과정이 저 사양 컴퓨팅 환경에 기계학습을 운용하기에 효과적임을 확인할 수 있었다.

5. 결론

본 논문에서는 저사양 컴퓨팅 환경에서 기계학습을 운용하기 위해 주성분 분석을 기반으로 차원을 축소하는 전처리 기법을 제안하였다. 또한 실험을 통해 기본적인 기계학습에서 정확도는 약 5% 미만으로 저하되었으나 성능은 96.6% 향상되었고 자원량은 99.7% 감소하는 것을 확인하였다. 따라서 제안하는 정책이 제한된 RAM과

프로세서를 가진 임베디드 환경, 엣지 컴퓨팅, 모바일 환경, 저개발 지역, 교육 및 연구 환경에서 기계학습을 응용하는데 효과가 있을 것으로 예상된다. 향후에는 저 사양 환경에서도 기계학습의 정확도를 더 향상할 수 있는 방법을 연구할 예정이다.

REFERENCES

- [1] F.Kherif and A.Latypova, "Principal component analysis," In Machine Learning, Elsevier, pp.209-225, 2020.
- [2] S.Karamizadeh, S.M.Abdullah, A.A. Manaf, M.Zamani and A.Hooman, "An overview of principal component analysis," Journal of Signal and Information Processing, Vol.4, pp.173-175, 2020.
- [3] A.C.Márquez, "The Curse of Dimensionality," In Digital Maintenance Management, pp.67-86, 2022.
- [4] B.Zhu, Z.Zhang, J.Dang, L.Wu and L.Wang, "Low-Complexity Visible Light Positioning and Rotation Estimation Based on Eigenvalue Decomposition," Journal of Lightwave Technology, Vol.20, No.21, pp.7072-7083, 2022.
- [5] C.Yang and Q.Shi, "An interval perturbation method for singular value decomposition (SVD) with unknown-but-bounded (UBB) parameters," Journal of Computational and Applied Mathematics, Vol.436, pp.115436, 2024.
- [6] X.Zhu, J.Li, Y.Liu, C.Ma and W.Wang, "A Survey on Model Compression for Large Language Models," Transactions of the Association for Computational Linguistics, Vol.12, pp.1556-1577, 2024.
- [7] B.Rokh, A.Azarpeyvand and A.Khanteymooi, "A Comprehensive Survey on Model Quantization for Deep Neural Networks in Image Classification," ACM Transactions on Intelligent Systems and Technology, Vol.14, No.6, pp.1-50, 2023.
- [8] B.Zhao, Q.Cui, R.Song, Y.Qiu and J.Liang, "Decoupled Knowledge Distillation," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.11953-11962, 2022.
- [9] https://www.kaggle.com/datasets/moltean/fruits/data?select=fruits-360_100x100
- [10] A.Geron, "Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow," O'Reilly Media, Inc, 2022.
- [11] S.Raschka, Y.Liu and V.Mirjalili, "Machine Learning with PyTorch and Scikit-Learn: Develop machine learning and deep learning models with Python," Packt Publishing Ltd, 2022.
- [12] M.K.Tran, S.Panchal, V.Chauhan, N.Brahmbhatt, A.Mevawalla, R.Fraser and M.Fowler, "Python-based scikit-learn machine learning models for thermal and electrical performance prediction of high-capacity

lithium-ion battery," International Journal of Energy Research, 2021.

- [13] F.Nelli, "Machine Learning with scikit-learn," In Python Data Analytics, pp.259-287, 2023.
- [14] E.C.Zabor, C.A.Reddy, R.D.Tendulkar and S.Patil, "Logistic Regression in Clinical Studies," International Journal of Radiation Oncology*Biology*Physics, Vol.112, No.2, pp.271-277, 2022.
- [15] S.Jin, C.Zhai, J.Liu, D.Jiang, D.Wang, S.Wang, and Y.Zhou, "Seismic fragility analysis of nuclear containment structure using Bayesian logistic regression model," Soil Dynamics and Earthquake Engineering, Vol.178, pp.108435, 2024.

이 현 섭(Hyun-Seob Lee)

[종신회원]



- 2013년 2월 : 한양대학교 컴퓨터 공학과 (공학 박사)
- 2012년 3월 ~ 2021년 2월 : 삼성전자 책임연구원
- 2021년 3월 ~ 현재 : 백석대학교 컴퓨터공학부 조교수

<관심분야>

인공지능, 저장시스템, 임베디드 시스템