



인공신경망과딥러닝입문

Lecture 04. 인공지능 개발을 위한 파이썬 라이브러리

동덕여자대학교
데이터사이언스 전공
권 범

목차

- ❖ 01. 인공지능 개발에 활용되는 라이브러리
- ❖ 02. 머신러닝 라이브러리 `scikit-learn`

01. 인공지능 개발에 활용되는 라이브러리

02. 머신러닝 라이브러리 `scikit-learn`

01. 인공지능 개발에 활용되는 라이브러리

❖ 두 가지 프로그래밍 환경

① Cloud 방식

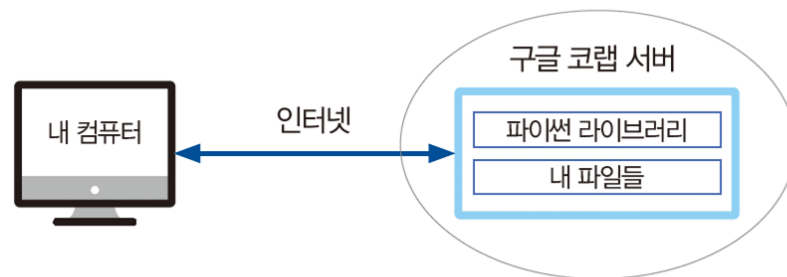
② Standalone 방식

정답은 없음.
나에게 적절한 환경을 선택하면 됨

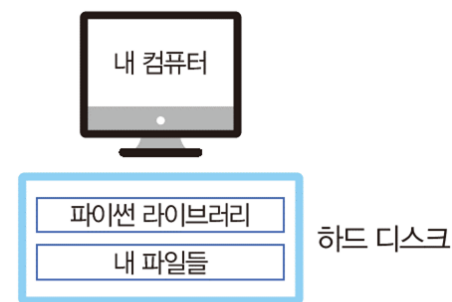
01. 인공지능 개발에 활용되는 라이브러리

❖ Cloud 방식

- 프로그램과 데이터가 서버에 저장되고 관리됨
 - 서버에 환경이 대부분 갖추어져 있어 로그인만 하면 바로 프로그래밍 가능함
 - 인터넷에 연결만 되어 있으면 어느 곳에서나 개발 및 협업이 가능함
 - 한계: 내 프로젝트에 최적의 환경을 갖추 수 없음
- ① Google: Colab(Colaboratory, 코랩)
 - ② Amazon: SageMaker(세이지메이커)
 - ③ Microsoft: Azure(애저)



(a) 클라우드 방식

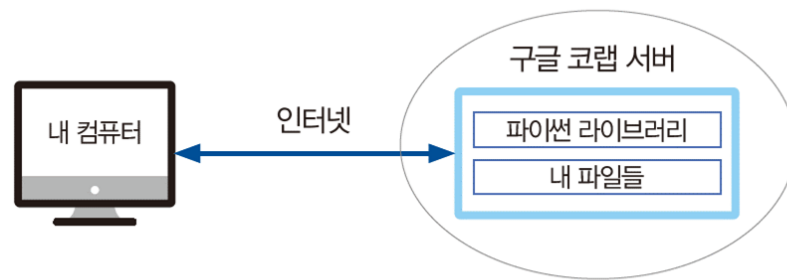


(b) 스탠드얼론 방식

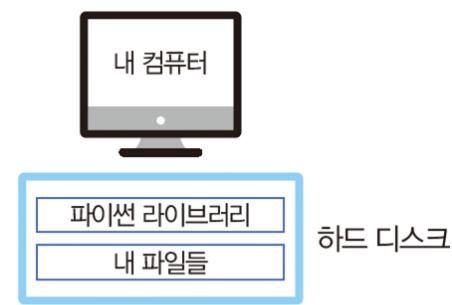
01. 인공지능 개발에 활용되는 라이브러리

❖ Standalone 방식

- 프로그램과 데이터가 자신의 컴퓨터에 저장됨
- 소프트웨어를 설치하고 환경을 스스로 구축해야 함
- 장점: 자신의 최적인 환경 구축이 가능함



(a) 클라우드 방식



(b) 스탠드얼론 방식

01. 인공지능 개발에 활용되는 라이브러리

❖ 프로그래밍 환경에 대한 좋은 태도 (1/2)

최근 여러 가지 프로그래밍 환경이 주어져 있음



인공지능 SW 프로그래머 층이 두터워졌다는 뜻으로 해석 할 수 있음

01. 인공지능 개발에 활용되는 라이브러리

❖ 프로그래밍 환경에 대한 좋은 태도 (2/2)

시범적으로 프로그래밍을 해보려는 사람.
본격적으로 제품을 개발하는 사람.
이 두 사람의 요구사항은 다를 수밖에 없음

따라서 다양한 프로그래밍 환경이
저마다의 목적으로 개발되고 있는 것임

두 가지 프로그래밍 환경 중에 반드시 하나를 선택해야 하는 것이 아니며,
두 환경 사이를 쉽게 전환할 수 있는 수준의 익숙함을 확보하는 것이
현명한 태도라고 볼 수 있겠음

01. 인공지능 개발에 활용되는 라이브러리

❖ Cloud 방식과 Standalone 방식의 선택 기준

- 간편한 실습 → Cloud 방식
 - ◆ 특히 단기 프로그래밍 과정에서는 설치 과정에 대한 시간을 절약할 수 있음
- 본격적인 개발 → Standalone 방식
 - ◆ 특히 인공지능 공학자로 성장하려고 하는 경우에는 스스로 최적의 환경을 구축할 수 있는 Standalone 방식이 적합함

01. 인공지능 개발에 활용되는 라이브러리

❖ 현대 프로그래밍 언어는 오픈소스 소프트웨어(Open-source Software, OSS)로 공개

- 파이썬(Python), R, Ruby, 펄(Perl), 줄리아(Julia), 스위프트(Swift) 등

◆ 특히 파이썬은 하루에도 수십 개의 새로운 라이브러리가 공개됨

무수히 많은 파이썬 라이브러리들 중에서
인공지능 개발에 활용되는 라이브러리에는 어떤 것들이 있을까요?

01. 인공지능 개발에 활용되는 라이브러리

❖ 인공지능 개발에 활용되는 라이브러리 종류

- 아래 표에 나와 있는 것들 이외에도 여러 라이브러리가 존재함
- SciPy(사이파이): 과학 계산용 함수(선형대수, 함수 최적화, 신호 처리 등)를 모아 놓은 라이브러리
- pandas(판다스): 데이터 처리와 분석을 위한 라이브러리

구분		설명	공식 사이트
프로그래밍 언어	Python(파이썬)		https://www.python.org
라이브러리 관리	PyPI(파이파이)	개발한 패키지를 업로드 & 공유	https://pypi.org
라이브러리	NumPy(넘파이)	다차원 배열 처리	https://numpy.org
	matplotlib(맷플롯립)	데이터 시각화	https://matplotlib.org
	scikit-learn(사이킷런)	기계학습 라이브러리	https://scikit-learn.org
	TensorFlow(텐서플로)	딥러닝 라이브러리	https://www.tensorflow.org
	Keras(케라스)	인공신경망 라이브러리	https://keras.io
	PyTorch(파이토치)	딥러닝 라이브러리	https://pytorch.org

02. 머신러닝 라이브러리 scikit-learn

01. 인공지능 개발에 활용되는 라이브러리

02. 머신러닝 라이브러리 scikit-learn

❖ scikit-learn이란?

- 파이썬 머신러닝 라이브러리
- 잘 알려진 머신러닝 알고리즘들은 물론 알고리즘을 설명한 풍부한 문서도 제공함
(<http://scikit-learn.org/stable/documentation>)
- 오픈소스인 scikit-learn은 자유롭게 사용 가능함
- 산업 현장이나 학계에도 널리 사용되고, 많은 튜토리얼과 예제 코드를 온라인에서 쉽게 찾을 수 있음
- scikit-learn은 다른 파이썬의 과학 패키지와의 잘 연동됨



02. 머신러닝 라이브러리 scikit-learn

❖ scikit-learn 설치 (1/4)

- scikit-learn은 NumPy와 SciPy 패키지를 사용하기 때문에, 이 두 패키지를 설치해야 함
- 또한, 그래프를 그리려면 matplotlib 패키지를,
대화식으로 개발하려면 IPython과 Jupyter Lab도 설치해야 함
- 필요한 라이브러리들을 모아 놓은 파이썬 배포판을 설치하는 것을 권장함
 - ① Anaconda: 대용량 데이터 처리, 예측 분석, 과학 계산을 위한 파이썬 배포판
 - ② Enthought Canopy: 과학 계산을 위한 파이썬 배포판
 - ③ Python(x,y): 윈도우 환경을 위한 과학 계산을 위한 무료 파이썬 배포판

02. 머신러닝 라이브러리 scikit-learn

❖ scikit-learn 설치 (2/4)

- scikit-learn의 패키지 의존성(Package Dependency) (2023-01-25일자 기준)

Dependency	Minimum Version	Purpose
numpy	1.17.3	build, install
scipy	1.3.2	build, install
joblib	1.1.1	install
threadpoolctl	2.0.0	install
cython	0.29.24	build
matplotlib	3.1.3	benchmark, docs, examples, tests
scikit-image	0.16.2	docs, examples, tests
pandas	1.0.5	benchmark, docs, examples, tests
seaborn	0.9.0	docs, examples
memory_profiler	0.57.0	benchmark, docs
pytest	5.3.1	tests
pytest-cov	2.9.0	tests
flake8	3.8.2	tests
black	22.3.0	tests
mypy	0.961	tests
pyamg	4.0.0	tests
sphinx	4.0.1	docs
sphinx-gallery	0.7.0	docs
numpydoc	1.2.0	docs, tests
Pillow	7.1.2	docs
pooch	1.6.0	docs, examples, tests
sphinx-prompt	1.3.0	docs
sphinxext-opengraph	0.4.2	docs
plotly	5.10.0	docs, examples
conda-lock	1.3.0	maintenance

[사진출처] <https://scikit-learn.org/stable/install.html>

02. 머신러닝 라이브러리 scikit-learn

❖ scikit-learn 설치 (3/4)

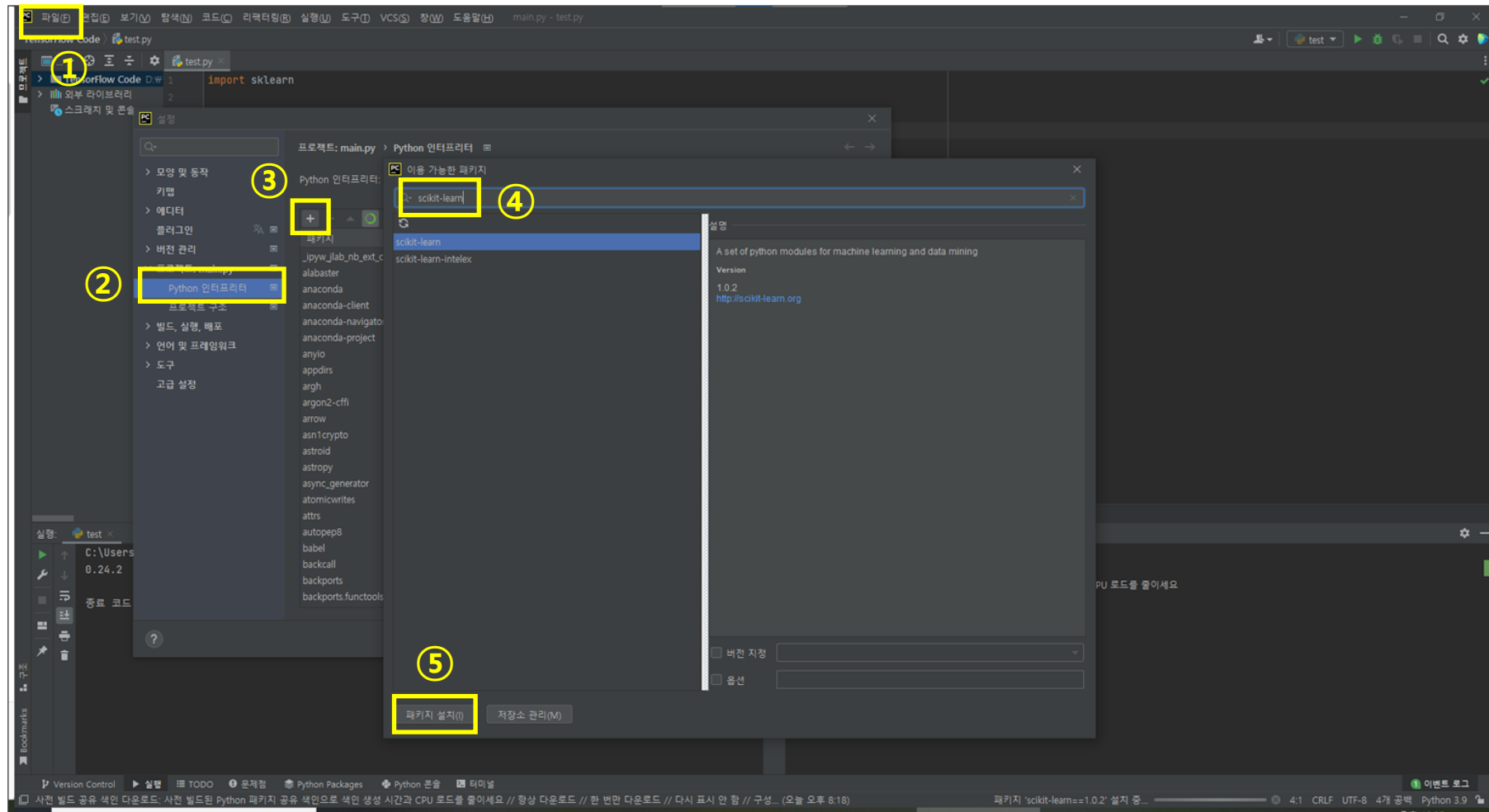
만약, Anaconda를 설치했다면
scikit-learn을 별도로 설치하실 필요가 없음

- ✓ scikit-learn 라이브러리 설치 명령어: `pip install scikit-learn`
- ✓ PyCharm 사용자: [설정] – Python 인터프리터 – scikit-learn 검색 후 패키지 설치

02. 머신러닝 라이브러리 scikit-learn

❖ scikit-learn 설치 (4/4)

- PyCharm 사용자: [파일]→[설정]→Python 인터프리터→[+] 버튼→scikit-learn 검색 후 패키지 설치



02. 머신러닝 라이브러리 scikit-learn

❖ scikit-learn 버전 확인

- scikit-learn 라이브러리를 불러오는 명령어: import sklearn

```
import sklearn  
  
print(sklearn.__version__)
```

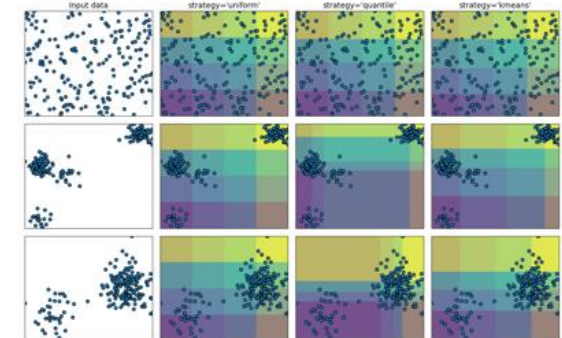
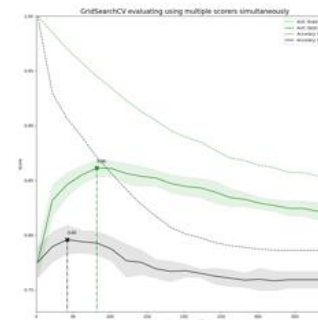
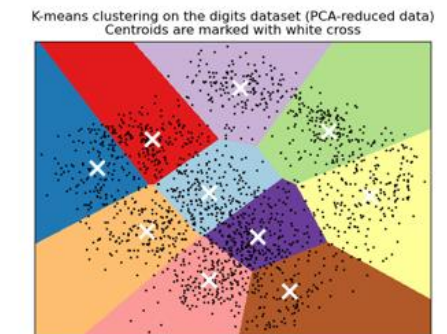
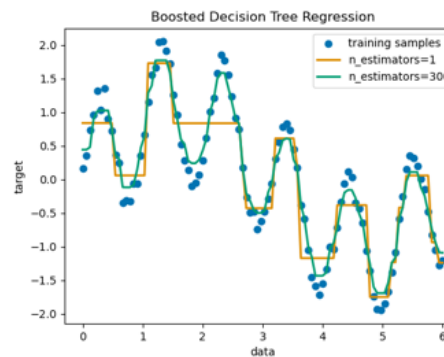
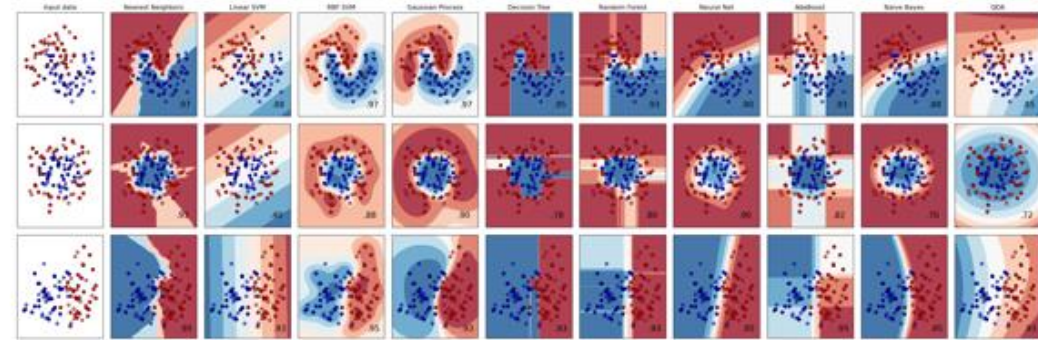
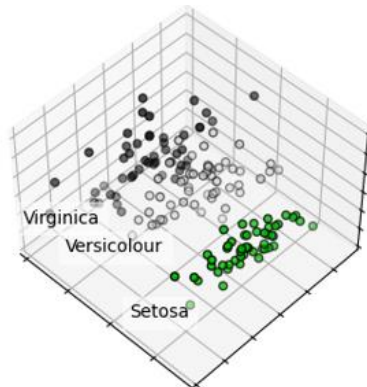
02. 머신러닝 라이브러리 scikit-learn

❖ scikit-learn으로 무엇을 할 수 있나요?

- 지도 학습
 - Classification 분류
 - Regression 회귀
- 비지도 학습
 - Clustering 군집화
 - Dimensionality Reduction 차원 축소 PCA(주성분 분석)

- Model Selection → `train_test_split()`
- Preprocessing 전처리
- etc.

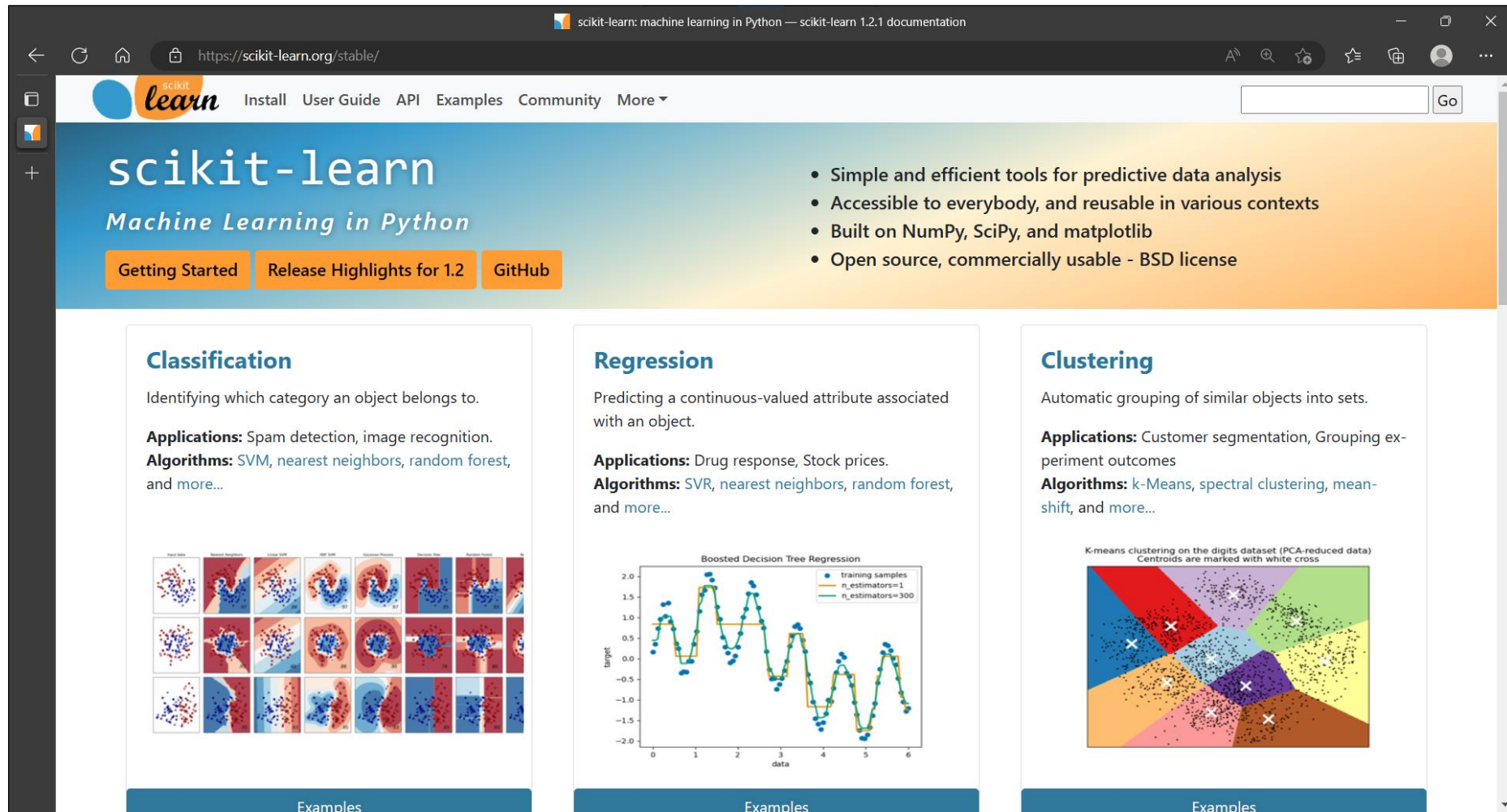
정규화, 표준화



02. 머신러닝 라이브러리 scikit-learn

❖ scikit-learn 공식 홈페이지

- [URL] <https://scikit-learn.org/stable/>



The screenshot displays the official scikit-learn website. At the top, there's a navigation bar with links for 'Install', 'User Guide', 'API', 'Examples', 'Community', and 'More'. The main header features the 'scikit-learn' logo and the tagline 'Machine Learning in Python'. Below this, there are three buttons: 'Getting Started', 'Release Highlights for 1.2', and 'GitHub'. To the right, a list of key features is provided:

- Simple and efficient tools for predictive data analysis
- Accessible to everybody, and reusable in various contexts
- Built on NumPy, SciPy, and matplotlib
- Open source, commercially usable - BSD license

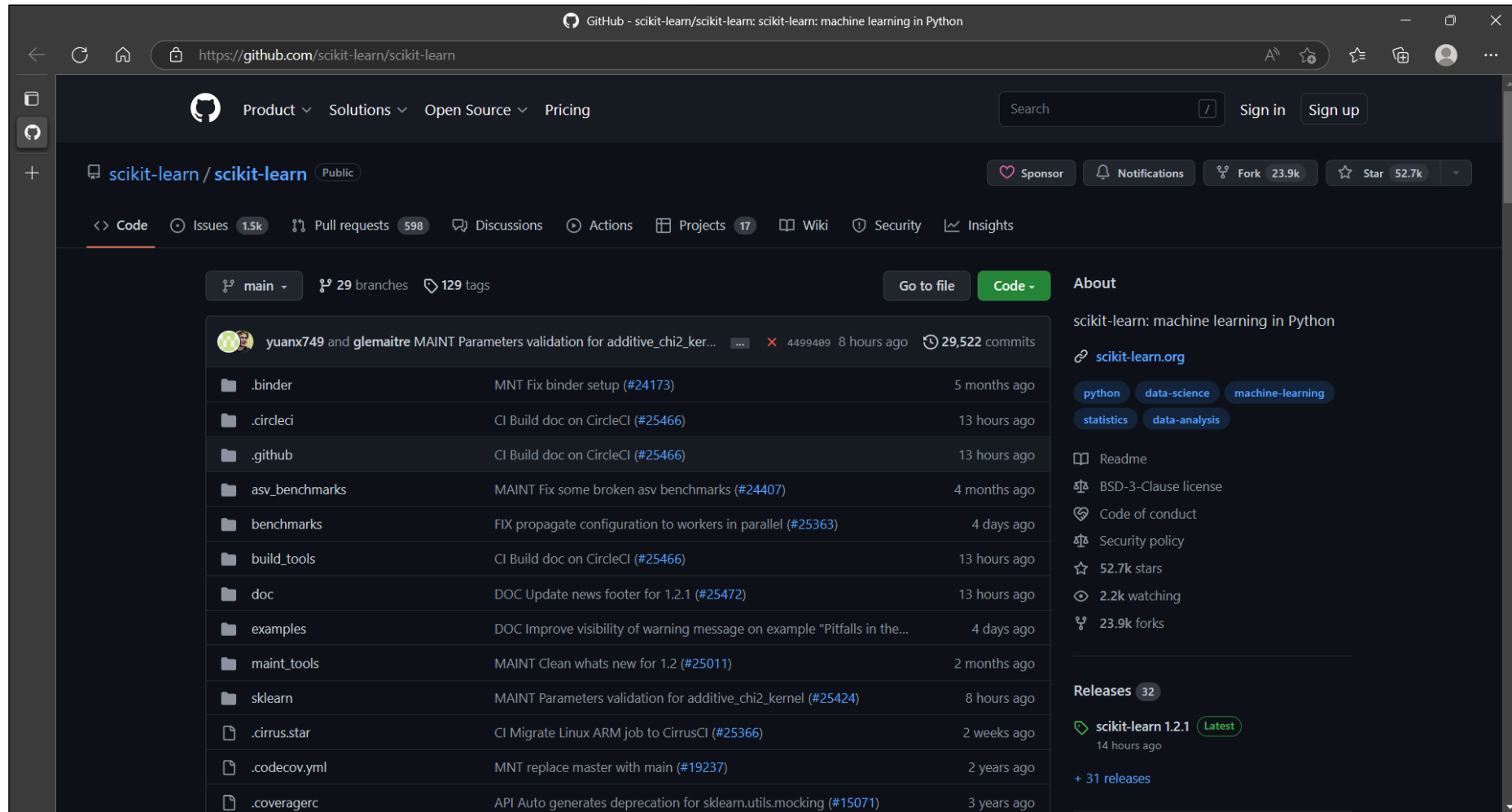
The page is divided into three main sections, each with a title, a brief description, applications, algorithms, and a visual example:

- Classification**: Identifying which category an object belongs to. Applications: Spam detection, image recognition. Algorithms: SVM, nearest neighbors, random forest, and more... The example shows a 3x3 grid of scatter plots with decision boundaries.
- Regression**: Predicting a continuous-valued attribute associated with an object. Applications: Drug response, Stock prices. Algorithms: SVR, nearest neighbors, random forest, and more... The example is a line plot titled 'Boosted Decision Tree Regression' showing 'target' vs 'data' with training samples and two regression lines (n_estimators=1 and n_estimators=300).
- Clustering**: Automatic grouping of similar objects into sets. Applications: Customer segmentation, Grouping experiment outcomes. Algorithms: k-Means, spectral clustering, mean-shift, and more... The example is a scatter plot titled 'K-means clustering on the digits dataset (PCA-reduced data)' showing data points grouped into clusters with centroids marked by white crosses.

02. 머신러닝 라이브러리 scikit-learn

❖ scikit-learn 공식 GitHub (1/2)

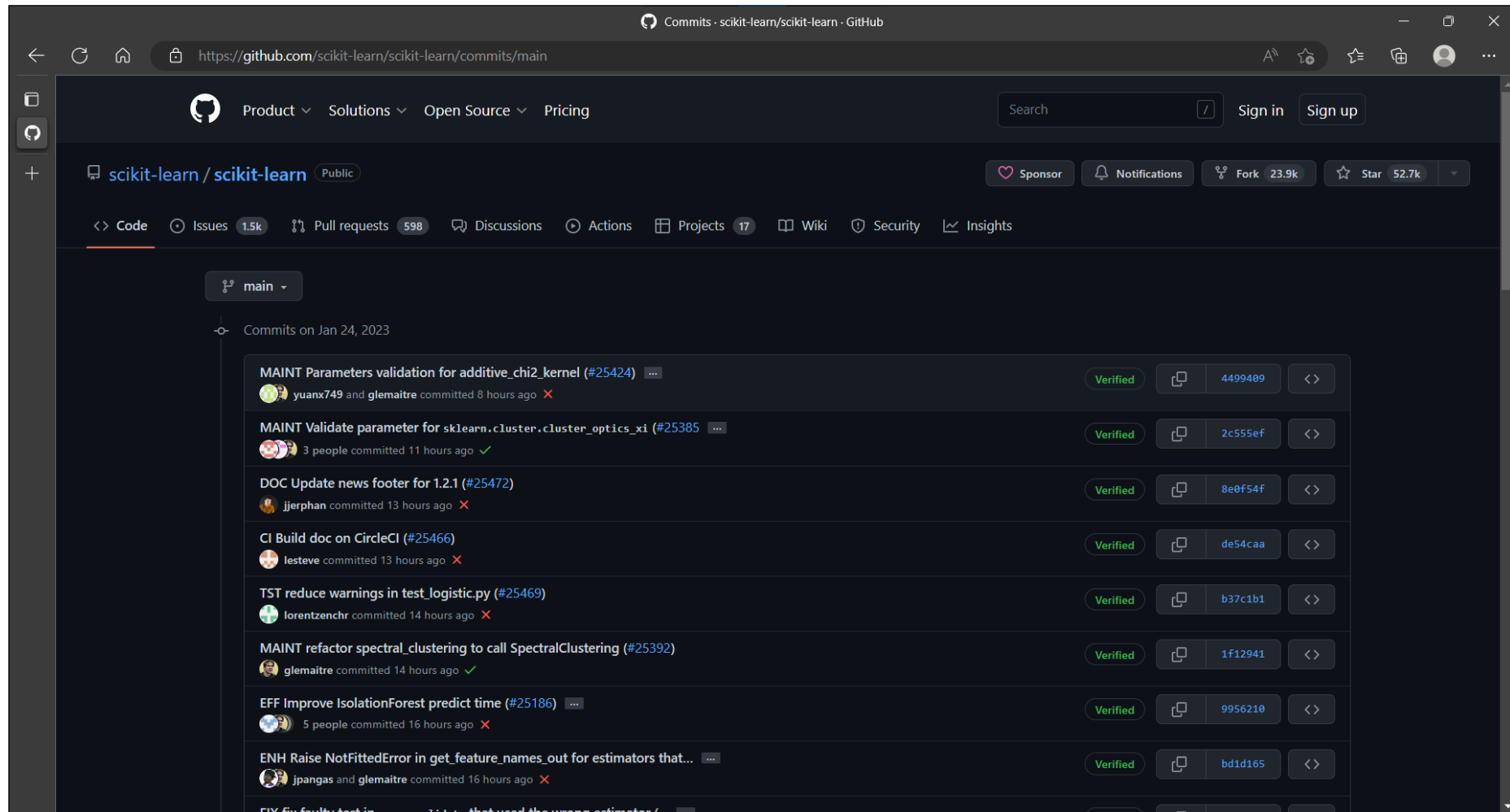
- [URL] <https://github.com/scikit-learn/scikit-learn>



02. 머신러닝 라이브러리 scikit-learn

❖ scikit-learn 공식 GitHub (2/2)

- 커밋 이력(Commit History)



- ❖ 01. 인공지능 개발에 활용되는 라이브러리
- ❖ 02. 머신러닝 라이브러리 `scikit-learn`

THANK YOU!

Q & A

- Name: 권범
- Office: 동덕여자대학교 인문관 B821호
- Phone: 02-940-4752
- E-mail: bkwon@dongduk.ac.kr