

머신러닝과 이상탐지

# 문제 중심 실습

강진범

# 문제 링크

<https://bit.ly/jb-ml-exam>

 EXAM.ipynb ☆

파일 수정 보기 삽입 런타임 도구 도움말 10월 27일에 마지막으로 수정됨

댓글 공유

+ 코드 + 텍스트

↑ ↓ ↶ ↷ ↵ ↶ ↷ ↵ ↶ ↷ ↵

### 프로젝트 제약사항

- 인터넷을 통해 API 문서를 살펴보고, 문제를 풀어보세요!!
- random seed=42로 설정합니다. (random\_state=42)

1. 97% 정확도(Accuracy)가 넘는 MNIST 분류기를 생성하세요.

- MNIST 데이터를 불러오세요.

```
from sklearn.datasets import fetch_openml
X_mnist, y_mnist = fetch_openml('mnist_784', return_X_y=True, as_frame=False)
```

- 데이터를 훈련 세트 60,000개, 테스트 세트 10,000개로 나누세요.

```
X_train, X_test, y_train, y_test = X[:60000], X[60000:], y[:60000], y[60000:]
```

- KNeighborsClassifier 분류기를 사용하세요.
- 그리드 서치(GridSearchCV)를 통해 파라미터를 최적화 하세요.
- 최적의 정확도를 가지는 분류기의 파라미터와 스코어를 출력하세요.
- (X\_train, y\_train)을 학습하고 (X\_test, y\_test)에 대한 score() 결과가 0.97보다 큼을 보이세요.

[ ]

2. 유사한 수준의 LinearSVC, SVC SGDClassifier 분류기 만들어 보세요.

# 사본 만들기

The screenshot shows the EXAM.ipynb interface. The 'File' menu is open, and the 'Save to Drive' option is highlighted with a red box. A red arrow points to this option with the text 'CLICK HERE'. The menu also includes options like 'Drive에 사본 저장', 'GitHub Gist로 사본 저장', 'GitHub에 사본 저장', '저장', '버전 저장 및 고정', '업데이트 기록', '다운로드', and '인쇄'.

EXAM.ipynb ☆

파일 수정 보기 삽입 런타임 도구 도움말 10월 27일에 마지막으로 수정됨

드라이브에서 찾기  
실습 모드에서 열기  
새 노트  
노트 열기 Ctrl+O  
노트 업로드  
이름 바꾸기  
이동  
휴지통으로 이동  
**Drive에 사본 저장** (CLICK HERE)  
GitHub Gist로 사본 저장  
GitHub에 사본 저장  
저장 Ctrl+S  
버전 저장 및 고정 Ctrl+M S  
업데이트 기록  
다운로드  
인쇄 Ctrl+P

에 의해 생산된 178개의 와인 샘플의 화학 분석 정보를 포함하고 있습니다. 이를 데이터에서 확인하세요.  
분리하세요. (test\_size=0.25, random\_state=42)  
자를 예측할 수 있는 분류 모델을 훈련 합니다.  
모든 3개의 클래스를 분류하기 위해 one-versus-all 방법을 사용해야 합니다.  
2가 넘는 SVM 분류기를 학습용 데이터셋으로 훈련하세요.

대한 SVM 회귀 모델을 훈련하고 RMSE 0.6보다 작은 모델로 파라미터 튜닝을 하세요

housing())을 사용하여 캘리포니아 주택 데이터 원본을 가져오세요.

- 데이터셋을 학습용, 테스트용으로 분리하세요. (test\_size=0.2, random\_state=42)
- 20,000개가 넘는 인스턴스가 있기 때문에 SVM은 느릴 수 있습니다. 따라서 하이퍼파라미터 튜닝을 위해 더 적은 2,000개 인스턴스를 사용 합니다.
- 테스트용 데이터셋에 대해서 RMSE가 0.6보다 작은 모델을 만드세요.

```
from sklearn.metrics import mean_squared_error  
rmse = mean_squared_error(y_test, y_pred, squared=False)
```

## 제약 사항

1. 제약사항 없이 자유롭게 문제를 풀어 봅니다.
2. 인터넷을 통해 API 문서를 살펴보고, 문제를 풀어보세요!!
3. 교재를 참고해 문제를 풀어보세요!!
4. 가능하면, 랜덤 random seed 42로 설정합니다. (random\_state=42)
5. 문제 가이드에 따라 퀴즈 문제를 풀어 봅니다.
6. 다양한 머신러닝 기술을 응용해 주어진 문제의 기대치를 높여 보세요!!

END