

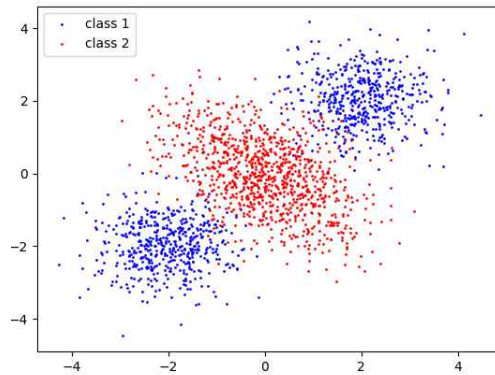
2019년도 1학기 패턴인식 Homework #3

- 개요

- 2D nonlinear 데이터를 생성하고, Multi-Layered Perceptron, RBF network, Kernel SVM, Random forest classifier에 대한 실험 수행
- 동일한 classifier를 이용하여 MNIST hand written digits 데이터 분류

- 요구사항

- classifier의 성능을 측정하기 위한 데이터를 생성한다.
- classification 데이터
 - ✓ 2 클래스로 이루어진 2차원 nonlinear 데이터를 생성한다.
 - ✓ 가우시안 분포를 이용하여 [그림 1]과 같은 형태로 데이터를 생성한다(평균과 분산은 자유롭게 설정).
 - ✓ 이 중 무작위로 70% 데이터를 학습 데이터로 사용하고, 나머지를 테스트 데이터로 사용한다.

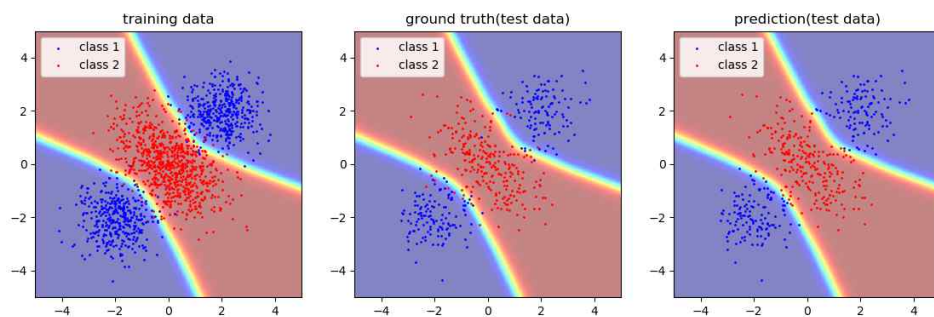


[그림 1] classification 데이터 예시

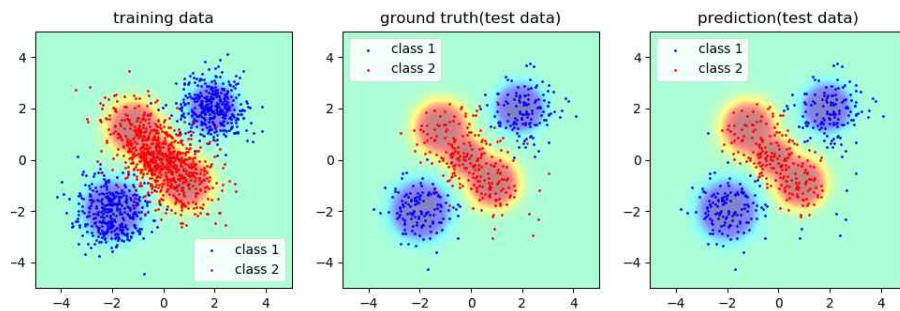
- 이 데이터에 대해 Multi-Layered Perceptron, RBF network, Kernel SVM, Random forest classifier을 이용하여 분류 성능을 측정한다.

- 예시

- Multi-Layered Perceptron

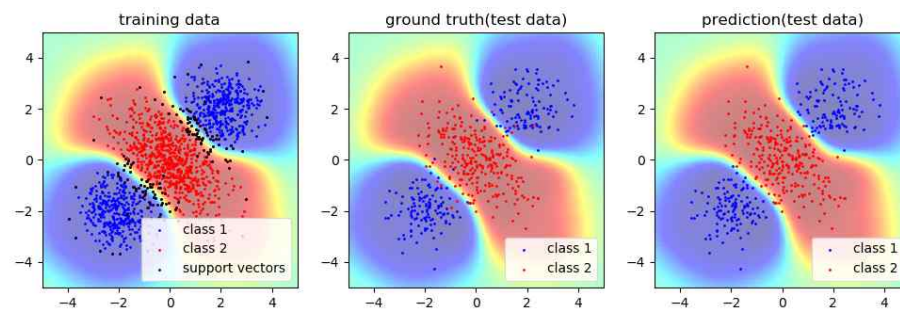


- RBF network



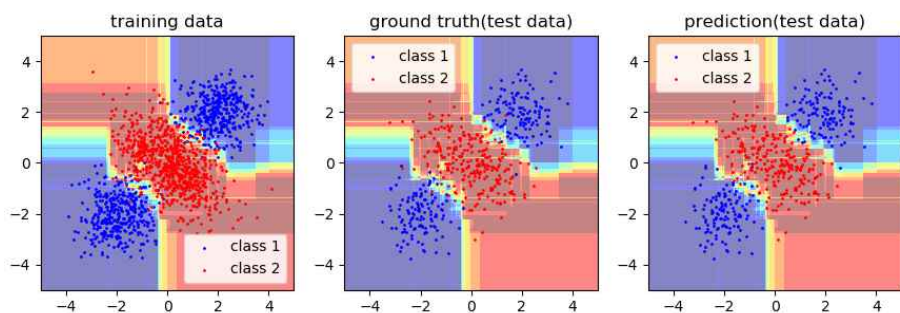
위 예제는 5개의 클러스터를 사용한 예제이다.
클러스터 수를 달리하며 실험한 결과를 보이시오.

- Kernel SVM



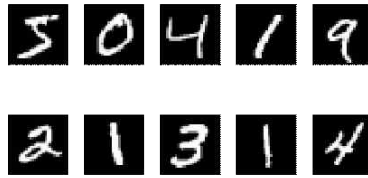
위 예제는 RBF를 kernel 함수로 사용한 예제이다.
kernel 함수를 달리하며 실험한 결과를 보이고, support vector를 표시하시오.

- Random forest



위 예제는 10개의 Decision tree를 사용한 예제이다.
Decision tree의 수를 달리하며 실험한 결과를 보이시오.

- MNIST hand written digits 데이터셋 분류



위 실험에서 사용한 classifier들을 이용하여 MNIST hand written digits 데이터를 분류한다. 또한 위 실험에서와 같이 각 classifier의 hyper-parameter나 kernel 함수를 달리하며 실험한다. 실험 결과로 confusion matrix를 제시하시오.

- **제출물**

- 요구사항을 구현한 소스 코드
- 플롯 결과를 포함한 한글 혹은 워드 문서
- 위 두 가지를 압축한 .zip 파일(파일명: 학번_이름_homework3.zip)

- **기타**

- Python 혹은 MATLAB 사용
- Python 사용 시, numpy, matplotlib, scikit-learn 사용 권장
- 이외의 어떠한 오픈소스 사용 가능