**K 최근접 이웃 분류 모델 미션(1)**

**K-최근접 이웃 분류 모델 구축 및 성능 분석**

오늘 배운 내용은 뒷 내용에도 계속 활용하며, 딥러닝을 진행할 때 역시 기본적으로 사용해야하는 기초적이면서도 가장 중요한 내용입니다. 잠깐 연습 해보는 시간을 가져보겠습니다. 아래 내용을 참고하여 가장 좋은 성능이 나오는 K 최근접 이웃 분류 모델을 만들어보세요. 성능 만큼 상대평가로 순위로 채점이 됩니다. 이 과정은 뒤에서 진행할 머신러닝 컴페티션 때 진행하는 과정에 매우 유사합니다.

[fish2.csv](attachment:7b4f795e-e64e-4bcb-bb31-b17f35aaa98a:fish2.csv)

**목표**

* **fish2 데이터셋**을 활용하여 K-최근접 이웃(KNN) 분류 모델을 구축하고, 다양한 파라미터(거리 계산 방법, 이웃 수)에 따른 모델의 성능을 분석합니다.

**과제 내용**

1. **모델 구축**
   * fish2 데이터셋을 사용하여 K-최근접 이웃 분류 모델을 구현하세요.
2. **데이터 분할**
   * 전체 데이터를 **훈련 데이터 : 시험 데이터 = 8 : 2**의 비율로 분할하세요.
3. **거리 계산 방법 평가**
   * 시험 데이터에 대해 **여러 가지 거리 계산 방법**(예: 유클리드, 맨해튼 등)을 적용해보세요.
   * 각 방법에 따른 모델의 성능(정확도)을 비교하여, **가장 좋은 성능을 보인 거리 계산 방법**을 선택하세요. **(데이터 표준화가 적용되어있어야 합니다.)**
4. **최적 이웃 수(k) 평가**
   * 시험 데이터를 대상으로 다양한 이웃 수(k)를 적용하여 모델을 평가하세요.
   * 가장 좋은 성능(정확도)을 보인 이웃 수(k)를 결정하세요.
5. **최적 모델 도출을 위한 시뮬레이션**
   * 각 **거리 계산 방법**별로 이웃 수를 **1부터 30까지** 변화시키며 시뮬레이션을 진행하세요.
   * 이 시뮬레이션 결과 중 **가장 높은 정확도**를 기록한 **데이터 표준화 방법**, **거리 계산 방법과 이웃 수**를 최종적으로 선정하세요.
6. **모델 성능 분석**
   * 5번에서 도출한 최적 모델을 이용하여 **물고기 종별 정밀도**를 비교하세요.
   * **가장 정밀도가 높은 물고기 종**과 **가장 정밀도가 낮은 물고기 종**이 무엇인지 확인하세요.

**제출 시 주의사항**

* **코드**와 함께 **분석 결과 리포트**를 제출하세요.
* 각 단계별 **진행 과정**과 **결과에 대한 해석**을 상세히 기술해야 합니다.
* 리포트는 곽경일 강사에게 Slack DM으로 제출하세요.