
Data Pattern Recognition – Term Project Guidelines

Instructor: Prof. Kyungho Won

Assignment Overview

현재 진행 중이거나 과거에 수행했던 연구 주제와 본 수업에서 다룬 내용을 연결하여 발표를 진행하고 보고서를 작성하시오

본 과제의 목적은 신호, 이미지, 센서, 텍스트 등의 실제 데이터에 ****전처리****, ****패턴 인식****, 또는 ****시각화 기법****을 어떻게 적용할 수 있는지를 보여주는 데 있음.

Expected Learning Outcome

- 데이터 패턴인식 수업에 나온 핵심 기법에 대한 이해를 보여줄 것
 - 전처리, 특징 추출, 차원 축소, 분류, 군집화 등 수업에서 배운 개념을 자신의 데이터, 가상의 데이터, 또는 공개 데이터에 적용해볼 것
 - 분석적 사고력과 의사소통 능력 (발표)를 연습할 것
-

1. Presentation (10 minutes)

**** 완성된 결과일 필요 없이, 수업 내용을 적용해본 과정과 논리적 연결이 중요**

- 본 강의에서 다룬 데이터 분석기법을 어떻게 적용해볼 수 있을 지 정리하고, 분석 시도 결과 또는 계획을 공유하는 것이 목적
- 현재 수행하고 있는, 또는 수행할 연구를 본 수업 내용과 연관지어 발표도 가능
- 현재 수행하고 있는 연구가 없고, 수행할 연구도 알아가는 중이라면 한가지 강의 내용을 주제로 잡고 공부해 가상의 데이터 혹은 공개된 예제 데이터에 적용해본 과정과 결과를 정리해서 발표해도 무관함 - 강의 중에 나온 실습 코드를 수정해서 활용할 수 있으나, 예제 코드와 100% 동일한 것 (단순 복사/붙여넣기)은 인정하지 않음.

구성요소: 배경 및 데이터 설명 + 코드 + 시각화 + 고찰

예시: sklearn의 iris 데이터셋을 불러와 PCA + SVM 적용해보기, EEG 예제 데이터를 불러와 간단한 PSD, (full FFT, welch's method, wavelet, multitaper) 시각화해보기, t-SNE를 써서 clustering 이 잘 되는 지 확인해보기

각각의 경우에 방법에 대해 설명하고, 어떤 결과가 나올지, 그리고 얻은 결과를 바탕으로 해석, 고찰 해보기 등의 내용필요.

✓ Structure (Recommended)

- Research Background
현재 수행하고 있는 연구 분야, 또는 데이터의 특성 (e.g., point cloud, sound, ...) 어떤 구조를 띄고 있는지, 시간/주파수 영역 분석을 주로 하는지, ...
- Pattern Recognition Link
수업 시간에 배운 어떤 내용을 적용해보거나 해볼 것 인지 소개
예시
 - EEG data → PSD + PCA + LDA
 - Image data → t-SNE + CNN
 - Time series → feature extraction + SVM
 - BCI → CSP + classifier comparison
- Results and Discussion
 - 성능 지표 (예: 정확도)나 시각화 (적절한 plot)
 - 고찰, 각각의 방법에 전처리 하기 전과 후를 비교하거나, 차원 축소법을 적용하기 전과 후의 정확도나 시각화 차이를 비교할 수 있음.
 - Data의 분포확인, ...

✓ Tips

- 10분 내외 발표
- 8-12 슬라이드 준비 (따로 제한 사항은 아님)
- 명확한 시각화와 텍스트 포함 (그림만 있거나, 글만 있는 슬라이드 지양, 그림은 의미 있는 그림 일 것 (flow chart 또는 실제 데이터를 시각화한 그림))
- 실제 구현 과정이 있으면 좋음

2. 📄 Final Report (Simplified Guidelines)

✓ 언어

- 한국어 또는 영어 중 자유롭게 선택

✓ 형식

- PDF 파일
- 글꼴/줄간격 자유 (A4 기준 꼭 찬 한 페이지, 2-3쪽 권장)

- Script를 첨부해 어떤 과정인지 설명
 - 도표, 그림 삽입 권장 - 분석과 관련된 것 (웹에서 가져온 상관없는 그림 제외)
 - 수식 포함은 자유
-

✅ 간단한 구성 (권장 섹션)

1. 주제 소개 (*Introduction*)
 - 본인이 다룬 데이터나 연구/과제 주제 간단 소개
 - 해당 주제를 선택한 이유
 2. 데이터 설명 (*Data Overview*)
 - 어떤 데이터를 사용했는지 (종류, 형태, 예시 등)
 - 전처리 과정이 있다면 간단히 서술
 3. 수업 내용과의 연결 (*Methods / Link to Course*)
 - 수업 중 배운 기법 중 어떤 것을 적용했는지
 - 예: PCA로 시각화, SVM으로 분류해봄, clustering 해봄 등
 - 직접 코드로 해봤다면 간단한 구현 방식 서술
 4. 분석 결과 또는 관찰 내용 (*Results / Observations*)
 - 시각화 결과, 비교 결과, 느낀 점 등
 - 결과가 없더라도 시도한 과정을 설명하면 충분함
 5. 맺음말 (*Conclusion*)
 - 어떤 점을 배웠는지, 적용에 어려웠던 점은 무엇인지
 - 향후 해보고 싶은 확장 방향이 있다면 간단히 적기
 6. 참고자료 (선택) (*References, if any*)
 - 논문, 웹자료 등을 참고했다면 명시 (형식 자유)
-

- 실제 연구가 없어도 괜찮음: 수업 내용 기반의 적용 시도/계획/해석이면 충분
- 성능이 높게 나오지 않거나, 제대로 분리가 잘 되지 않아도 괜찮음. 적용하려고 했던 과정 자체가 학습의 일부이나, 해석/고찰 필요
-

예시: IRIS 데이터에 PCA + SVM 적용

슬라이드 또는 리포트 구성 예시

- 어떤 데이터셋인지 설명 (클래스 개수, feature 개수, 샘플 개수 등)
- 분석 목적: 분류 성능 확인, 회귀 성능 확인, 또는 시각화 실험
- 간단한 분석 목표 설정
 - 전처리: 어떤 스케일을 이용하거나 정규화 했는지 등, outlier 탐색, ...
 - 사용기법: PCA 개념 소개
- 왜 PCA를 이용했는 지 설명
 - 시각화 결과: PCA를 이용해 차원을 고차원 -> 2차원으로 줄인 후 2D scatter plot 등에서 시각화하고, 클래스 별로 color mapping을 통해 분리가 되는 지 눈으로 확인
 - SVM 적용 결과 - decision boundary, 분류 정확도 등, PCA를 적용하기 전, 적용한 후의 분류 정확도 비교 등
 - 고찰: 성능이 어떤 지, 어떤 클래스가 구분되는지
 - 수업과의 연결
 - 한계점 및 향후 방향: iris dataset의 샘플 개수가 부족하거나 너무 단순해 PCA가 큰 변화를 내지 않았다, 다른 모델을 고려해볼 수 있다 등