

# Healthcare Application Software Final Project



**BDSL**  
Biomedical Data Science Lab.

## Task: 건강보험 데이터를 이용한 고혈압 예측

- 데이터
  - 국민건강보험공단 건강검진정보
  - 2022년: <https://www.data.go.kr/data/15007122/fileData.do>
- 라벨링
  - 고혈압: 수축기 혈압 140 이상이면서 (and) 이완기 혈압이 90 이상인 경우 1
  - 정상: elsewhere 0
  - Reference: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9548597/>
- 사용 기술: 헬스케어응용SW에서 배운 내용
  - 데이터 전처리: numpy, pandas 등
  - 그래프: matplotlib, seaborn 등
  - 예측 모델: scikit-learn
  - 프로젝트 페이지: github.io, html, css, ...

# 보고서 작성

- 프로젝트 페이지에 포함되어야 하는 내용
  - 서론
    - 건강검진 데이터로 혈압을 예측해야 하는 의미
    - 본인이 구현한 코드의 핵심 내용 요약
  - 본론
    - 데이터 전처리 과정
    - 사용한 인공지능 모델
    - 실험 결과
    - 실험 결과에 대한 설명
  - 결론
    - 보고서 전체 내용을 아우를 수 있는 문단
  - 참고문헌
    - 다른 자료 및 논문의 내용을 인용한 경우 반드시 인용한 문장 끝에 [1] 등으로 참고문헌 번호를 표기하고, 보고서 가장 마지막에 참고문헌을 [1] 부터 순서대로 나열하여야 함

# 평가항목 및 유의사항

- 평가항목
  - 데이터 전처리의 적절성
  - 예측 모델의 성능
  - 표, 그래프, 그림 등으로 실험 과정과 결과를 잘 표현하였는지
  - 프로젝트 페이지에 작성한 보고서
- 유의사항
  - 데이터가 너무 많은 경우, 데이터를 선별하여 사용하여도 됨. 다만 보고서에 이유와 선별 기준을 명시하여야 함
  - 데이터가 부족한 경우, 다른 연도의 건강검진정보를 이용하여도 됨. 마찬가지로 보고서에 명시하여야 함
  - 전처리 과정, 모델, 실험결과 시각화 모두 자유이며, 프로젝트 페이지의 보고서에 상세히 작성하여야 함
  - 마감 기한: 6/17 (월) 23:59
    - Late day를 고려하여 마감 기한에 맞추어 제출할 것

Thank You!