2024 SNUBH Datathon예선

예선 성능 평가 [NARAK]

주최/주관 의료인공지능센터 Center for Artificial Intelligence in Healthcare | SNU AI.MED 서울대학교 의료 인공지능 융합인재 양성 사업단 | AI NATION

후원 aws AWS코리아 | KHIDI 한국보건산업진흥원 | IITP 정보통신기획평가원

Contents

팀 기본정보

·팀명: NARAK(Nephrology Advanced Research and Analysis in Korea)

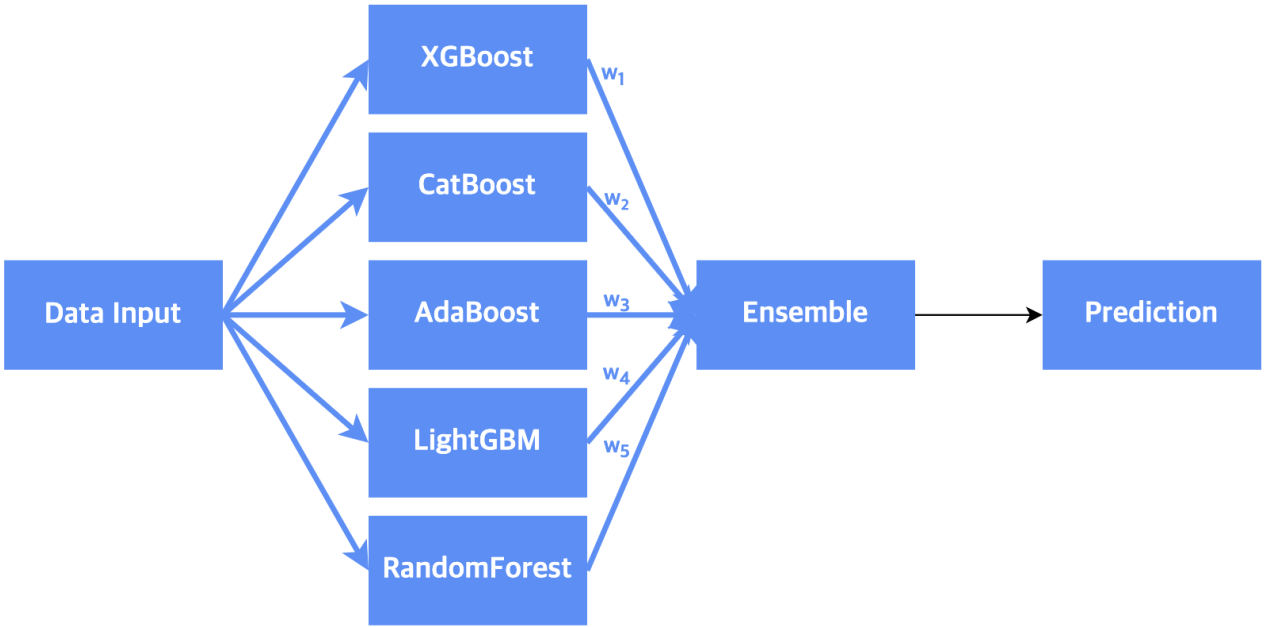
· 팀원 설명

- 박준범 (서울대학교 기계공학부 20학번)
- 김시훈 (서울대학교 전기정보공학부 21학번)
- 이윤서 (가톨릭대학교 컴퓨터공학부 20학번)
- 손재혁 (서울대학교 전기정보공학부 24학번)

AI 모델명

· AKI 여부를 예측하기 위해서 이진 분류 모델을 사용

· 기본 모델: XGBoost, CatBoost, AdaBoost, LightGBM, RandomForest 를 각 F1-score의 가중치를 사용하여 enser



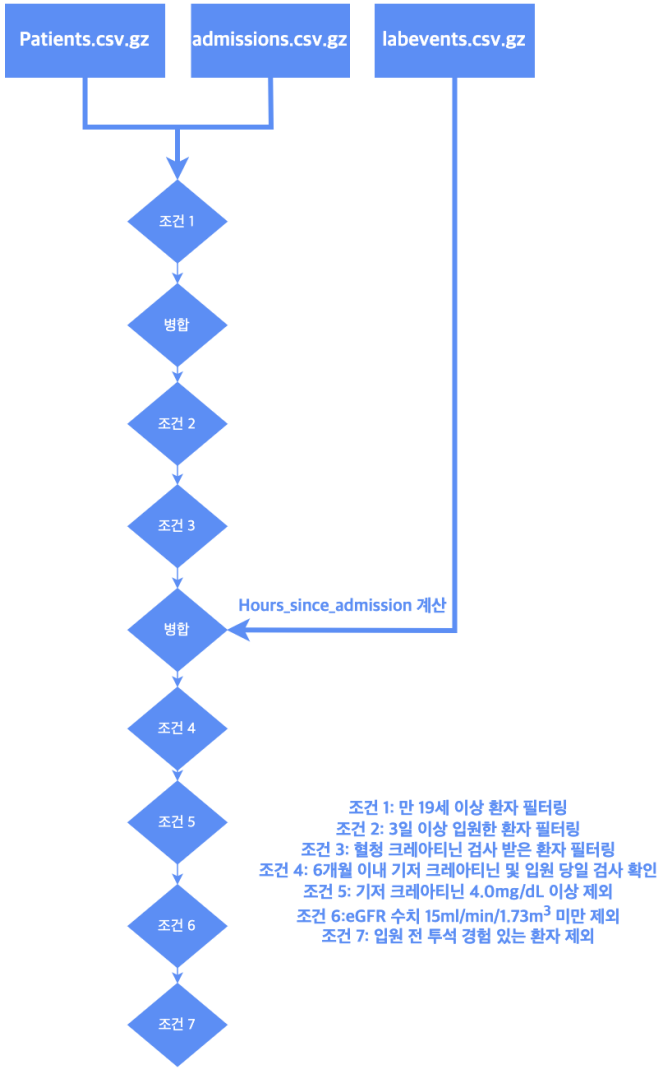
모델 상세정보

· Dataset의 피쳐들을 추출하여 조건들에 따라 필터링 후 병합

· 결측치가 발생 할 경우 결측치 비율 계산하며 결측치 제거

· 각 피처 정보들을 병합 및 정리 실행

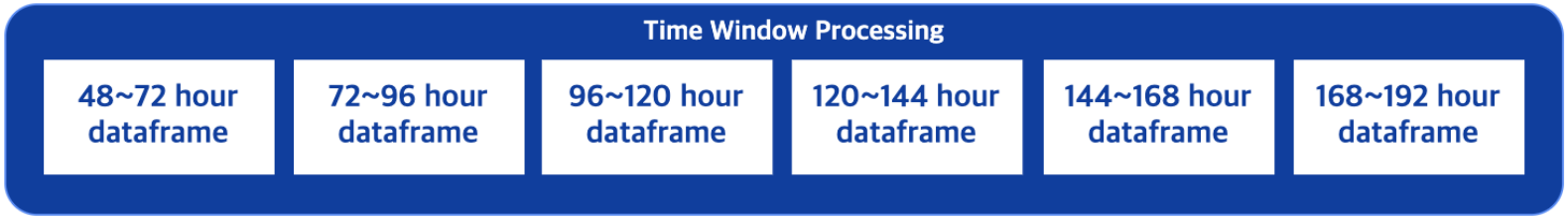
· 피처들을 Time Window로 나눠 트레이닝 및 결과 추출



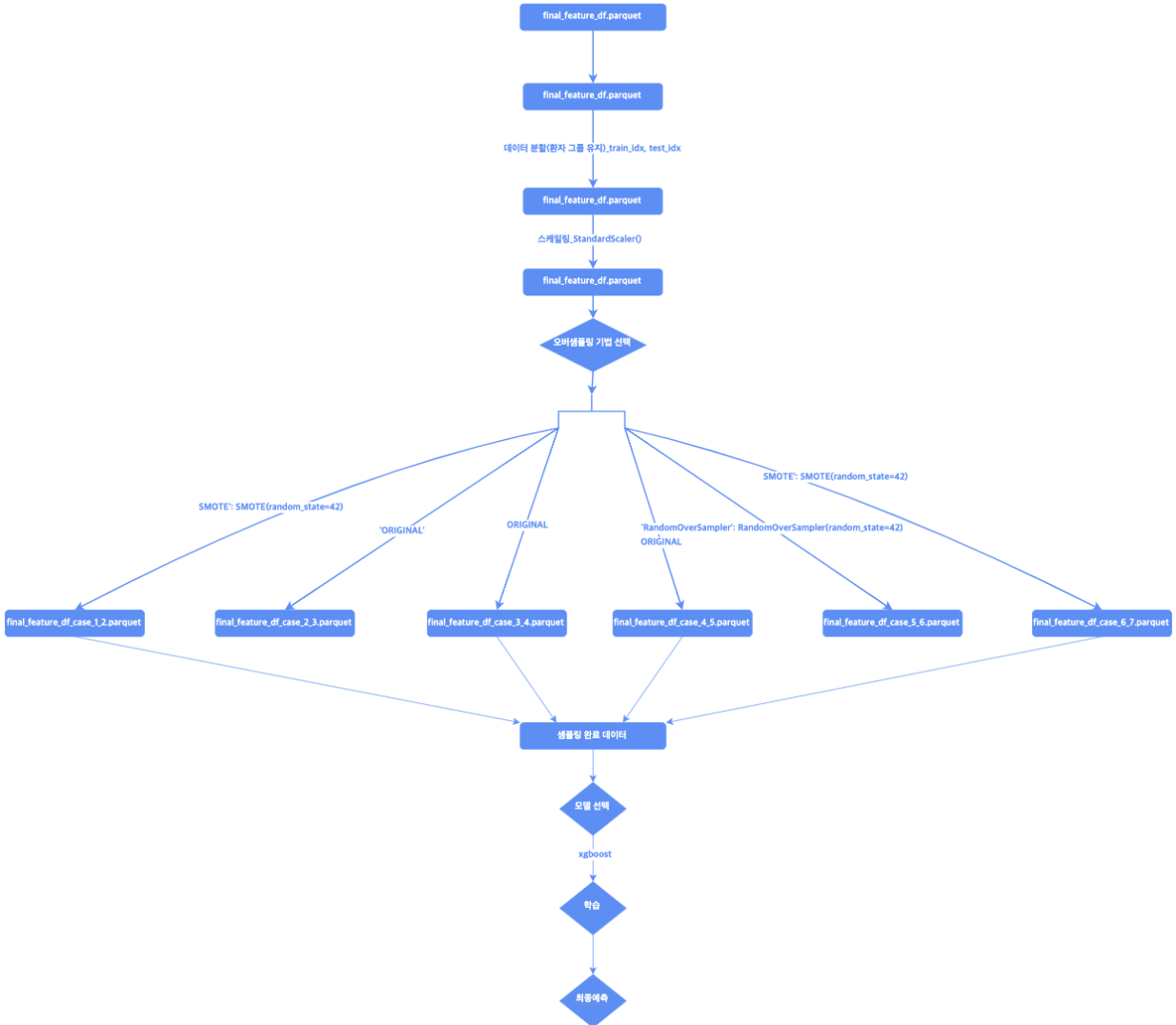
Extracted Features

- Age
- Gender
- Urine Output
- Vasopressor Amount
- Weight
- BMI
- Race
- Admission Time
- Length of Stay
- Baseline Creatinine
- eFGR
- Number of Diagnoses
- Number of Procedures
- Drug severity
- Drug mortality
- Usage of Nephrotoxic
- Renal Replacement Therapy
- Invasive mechanical ventilation
- Lab Measurements(Hemoglobin, Potassium etc.)
- Vital signs(Heart rate, Respiratory rate etc.)

데이터 정리 및 결측치 처리 & Time Window별 데이터셋 생성



학습

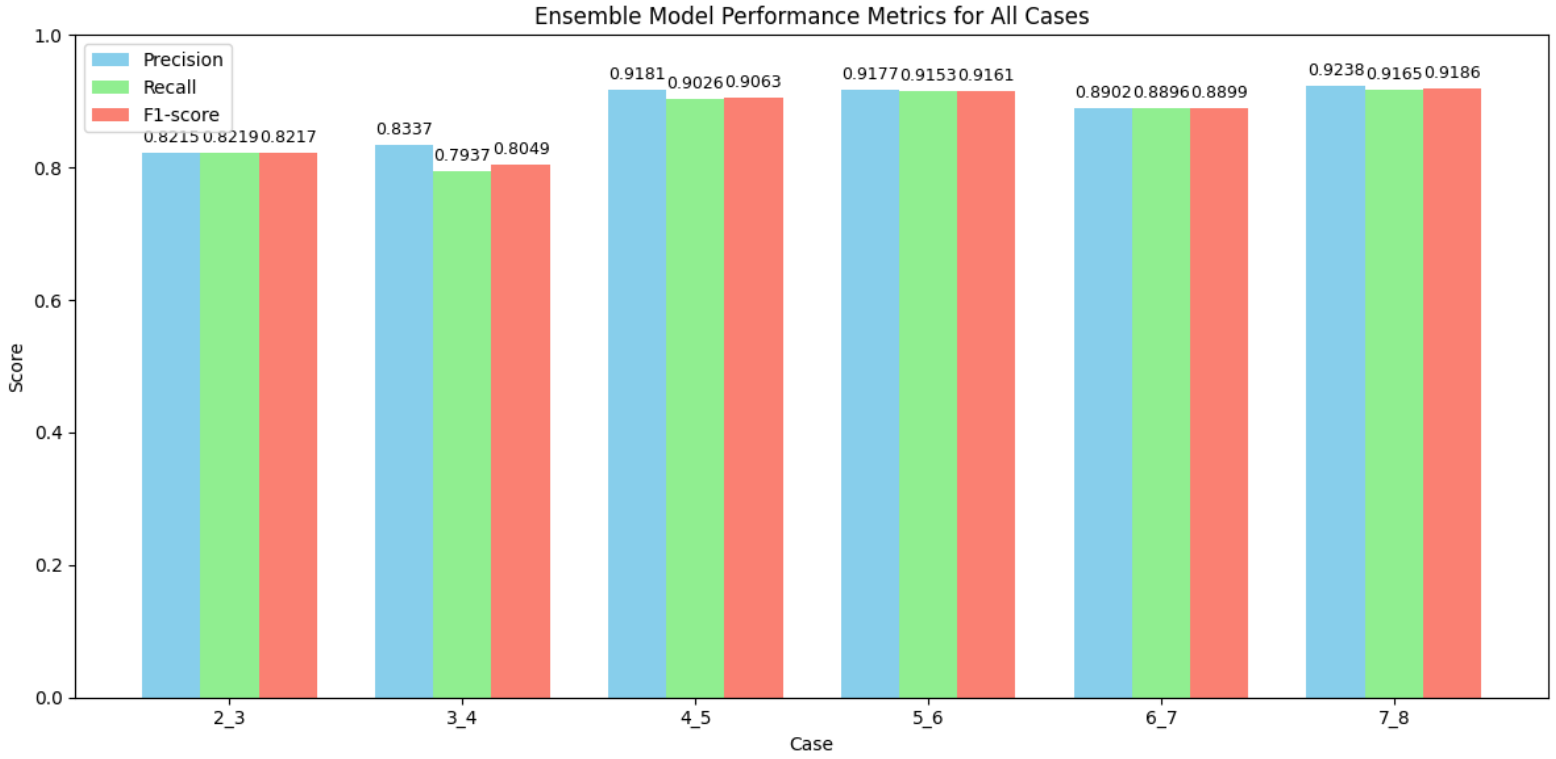

final_feature_df.parquet
final_feature_df.parquet
데이터 분할(환자 그룹 유지)_train_idx, test_idx
final_feature_df.parquet
스케일링_StandardScaler()
final_feature_df.parquet
오버샘플링 기법 선택
SMOTE': SMOTE(random_state=42)
'ORIGINAL'
ORIGINAL
'RandomOverSampler': RandomOverSampler(random_state=42)
ORIGINAL
SMOTE': SMOTE(random_state=42)
final_feature_df_case_1_2.parquet
final_feature_df_case_2_3.parquet
final_feature_df_case_3_4.parquet
final_feature_df_case_4_5.parquet
final_feature_df_case_5_6.parquet
final_feature_df_case_6_7.parquet
샘플링 완료 데이터
모델 선택
xgboost
학습
최종예측

개발환경

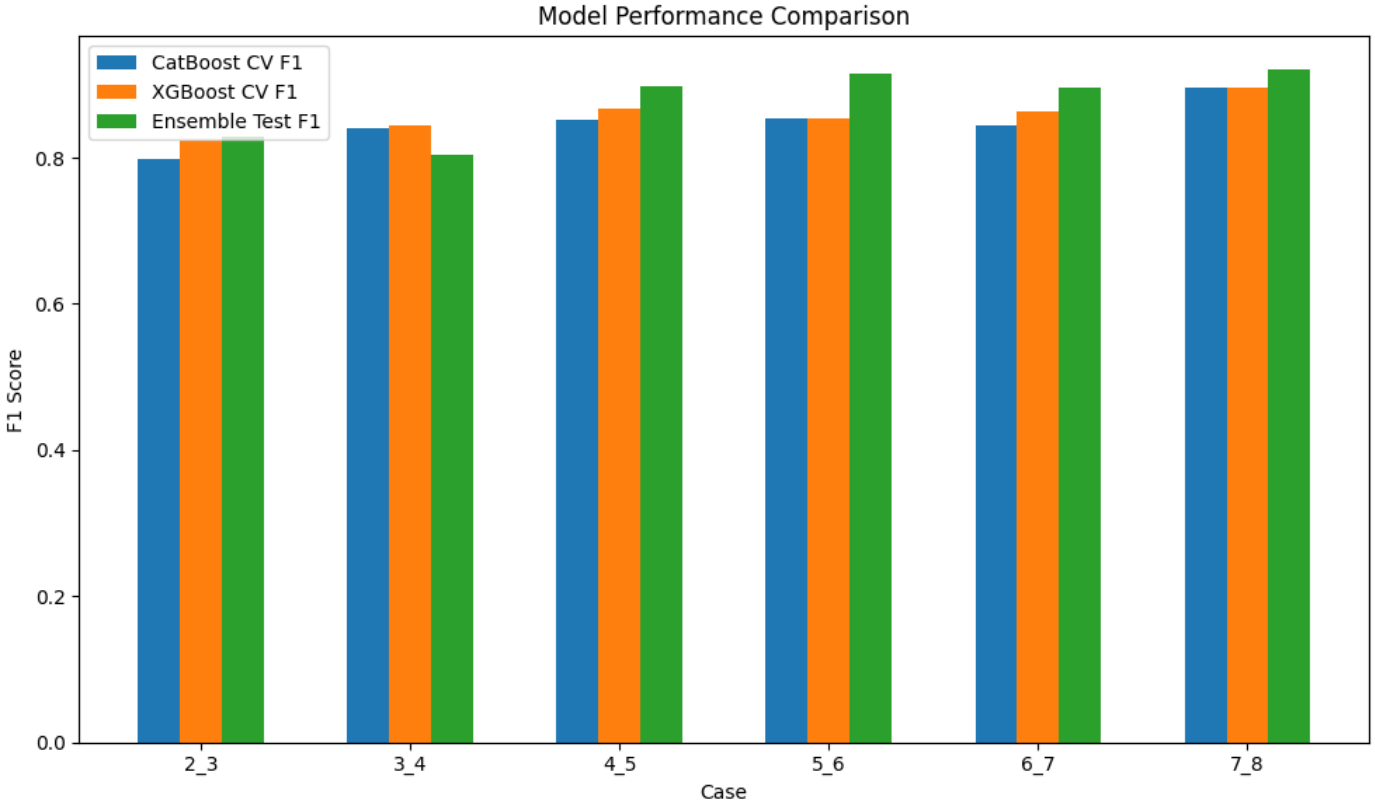
Package	Version
ace_tools	0.0
annotated-types	0.7.0
asttokens	2.4.1
backcall	0.2.0
blis	0.7.10
Brotli	1.1.0
catalogue	2.0.10
catboost	1.2.7
certifi	2024.8.30
cffi	1.17.0
charset-normalizer	3.3.2
clarabel	0.9.0
click	8.1.7
cloudpathlib	0.19.0
cloudpickle	3.0.0
colorama	0.4.6
comm	0.2.2
confection	0.1.4
contourpy	1.1.1
cramjam	2.8.3
cvxopt	1.3.2
cvxpy	1.5.2
cycler	0.12.1
cymem	2.0.8
dask	2023.5.0
debugpy	1.8.5
decorator	5.1.1
ecos	2.0.14
en-core-web-md	3.7.1
exceptiongroup	1.2.2
executing	2.1.0
fancyimpute	0.7.0
fastparquet	2024.2.0
filelock	3.16.1
fonttools	4.53.1
fsspec	2024.9.0
graphviz	0.20.3
h2	4.1.0
hpack	4.0.0
hyperframe	6.0.1
idna	3.10
imbalanced-learn	0.12.3
imblearn	0.0
importlib_metadata	8.5.0
importlib_resources	6.4.5
iniconfig	2.0.0
ipykernel	6.29.5
ipython	8.12.2
jedi	0.19.1
Jinja2	3.1.4
joblib	1.4.2
jupyter_client	8.6.3
jupyter_core	5.7.2
kiwisolver	1.4.7
knnimpute	0.1.0
langcodes	3.4.0
language_data	1.2.0
lightgbm	4.5.0
locket	1.0.0
marisa-trie	1.1.0
markdown-it-py	3.0.0
MarkupSafe	2.1.5
matplotlib	3.7.5
matplotlib-inline	0.1.7
mdurl	0.1.2
mpmath	1.3.0
murmurhash	1.0.10
nest_asyncio	1.6.0
networkx	3.1
nose	1.3.7
numpy	1.24.4
osqp	0.6.7.post1
packaging	24.1
pandas	2.0.3
parso	0.8.4
partd	1.4.1
pickleshare	0.7.5
pillow	10.4.0
pip	24.2
platformdirs	4.3.6
plotly	5.24.1
pluggy	1.5.0
preshed	3.0.9
prompt_toolkit	3.0.47
psutil	6.0.0
pure_eval	0.2.3
pyarrow	17.0.0
pycparser	2.22
pydantic	2.8.2
pydantic_core	2.20.1
Pygments	2.18.0
pyparsing	3.1.4
PySocks	1.7.1
pytest	8.3.3
python-dateutil	2.9.0
pytz	2024.2
pywin32	306
PyYAML	6.0.2
pyzmq	26.2.0
qdldl	0.1.7.post4
requests	2.32.3
rich	13.8.1
scikit-learn	1.3.2
scipy	1.10.1
scs	3.2.7
seaborn	0.13.2
setuptools	75.1.0
shellingham	1.5.4
six	1.16.0
smart_open	7.0.4
spacy	3.7.5
spacy-legacy	3.0.12
spacy-loggers	1.0.5
srsly	2.4.8
stack-data	0.6.2
sympy	1.13.3
tenacity	9.0.0
thinc	8.2.5
threadpoolctl	3.5.0
tomli	2.0.1
toolz	0.12.1
torch	2.4.1+cu118
torchaudio	2.4.1+cu118
torchvision	0.19.1+cu118
tornado	6.4.1
tqdm	4.66.5
traitlets	5.14.3
typer	0.12.5
typer-slim	0.12.5
typing_extensions	4.12.2
tzdata	2024.1
urllib3	2.2.3
wasabi	1.1.2
wcwidth	0.2.13
weasel	0.4.1
wheel	0.44.0
win-inet-pton	1.1.0
wrapt	1.16.0
xgboost	2.1.1
zipp	3.20.2
zstandard	0.23.0

실행방법

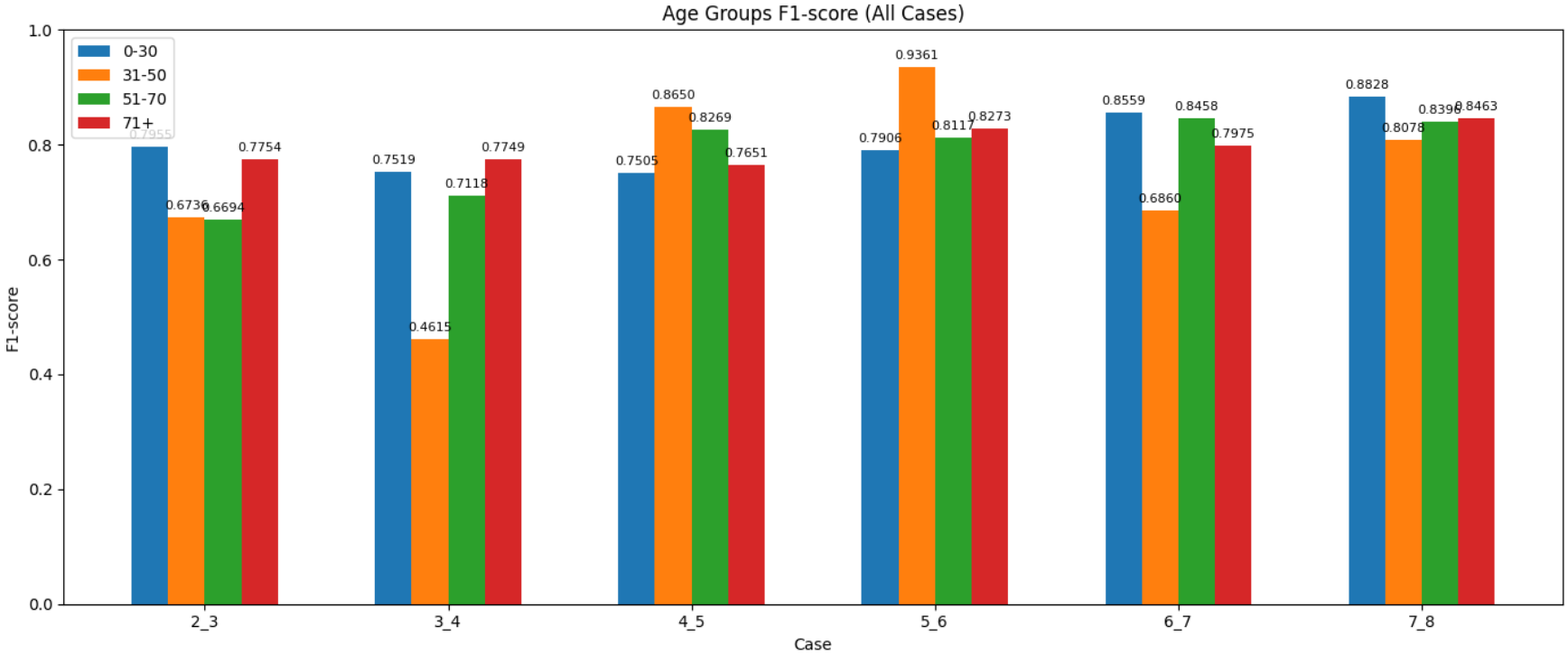
성능평가


Ensemble Model Performance Metrics for All Cases
Precision
Recall
F1-score
Score
Case
1.0
0.8
0.6
0.4
0.2
0.0
2_3
3_4
4_5
5_6
6_7
7_8
0.8215
0.8219
0.8217
0.8337
0.7937
0.8049
0.9181
0.9026
0.9063
0.9177
0.9153
0.9161
0.8902
0.8896
0.8899
0.9238
0.9165
0.9186

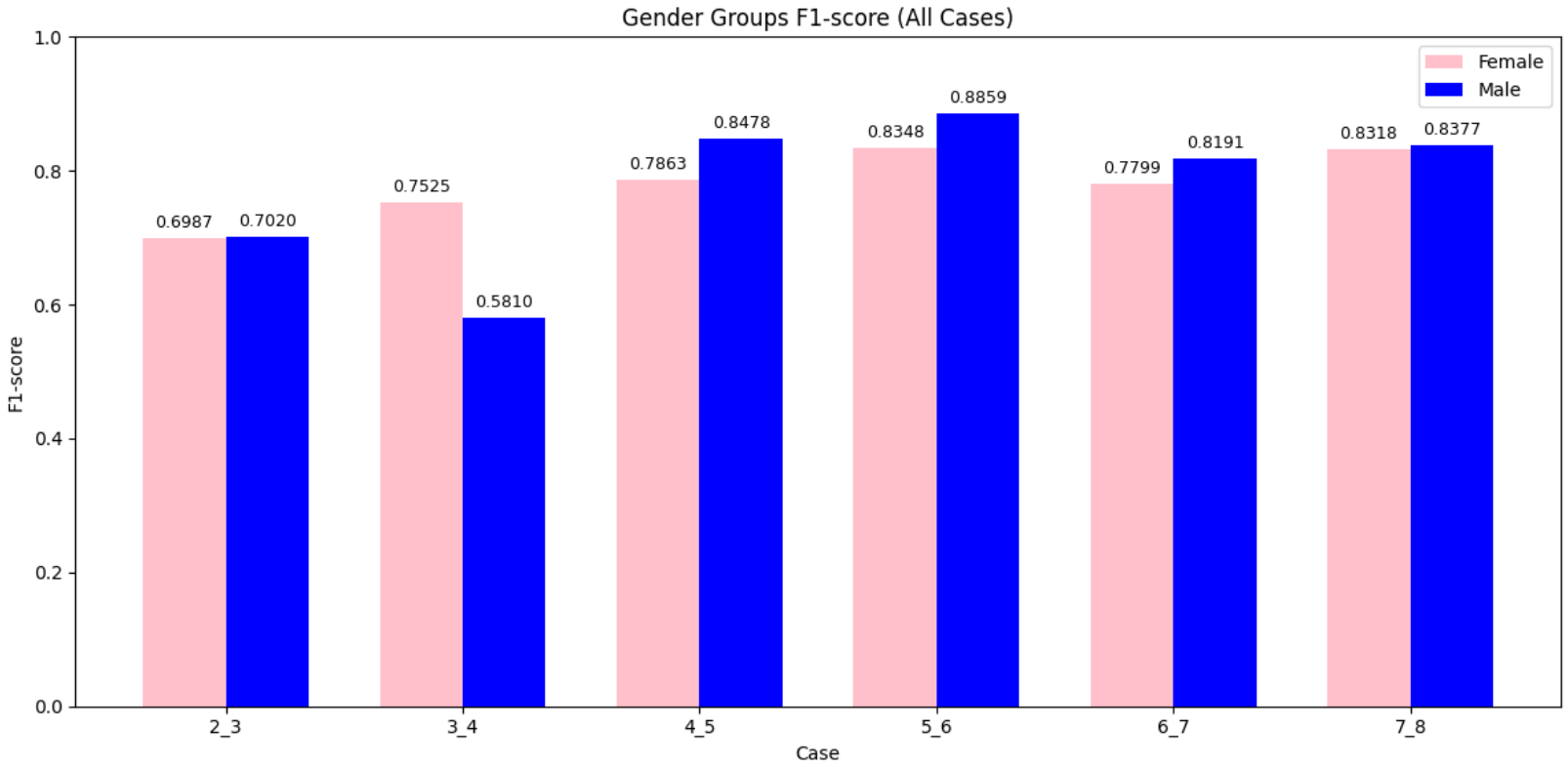
성능평가


Model Performance Comparison
CatBoost CV F1
XGBoost CV F1
Ensemble Test F1
F1 Score
0.0
0.2
0.4
0.6
0.8
2_3
3_4
4_5
5_6
6_7
7_8
Case

공정성평가


Age Groups F1-score (All Cases)
0-30
31-50
51-70
71+
F1-score
Case
1.0
0.8
0.6
0.4
0.2
0.0
2_3
3_4
4_5
5_6
6_7
7_8
0.7955
0.6736
0.6694
0.7754
0.7519
0.4615
0.7118
0.7749
0.7505
0.8650
0.8269
0.7651
0.7906
0.9361
0.8117
0.8273
0.8559
0.6860
0.8458
0.7975
0.8828
0.8078
0.8398
0.8463

공정성평가


Gender Groups F1-score (All Cases)
Female
Male
F1-score
Case
1.0
0.8
0.6
0.4
0.2
0.0
2_3
3_4
4_5
5_6
6_7
7_8
0.6987
0.7020
0.7525
0.5810
0.7863
0.8478
0.8348
0.8859
0.7799
0.8191
0.8318
0.8377

본선모델개발방법론

·MIMIC-IV 데이터를 전처리한 방법과 동일하게 서울대 병원 데이터들을 전처리

· 결측값, 이상치, 일관성 없는 데이터 항목 처리 방식 통일

· 대상 데이터셋에 대한 성능을 최적화하기 위해 전이 학습 프로세스 반복 조정

·또한 MIMIC-IV 데이터들로 계산한 pretrained weight들을 가져와 본선에 사용

·현재 예선 모델은 회귀만 사용하고 있으나 시계열 특성들을 따로 추출해 시계열 모델 적용

· 추세, 계절성, 주기적 패턴과 같은 시간 기반 특성 추출

·환자들의 종교, 윤리적 가치관, 혼인여부, 직업 등의 정보들이 존재할 시 모델에 고려

참고문헌

1. Choi, Y. H., Kim, S. H., Ahn, S. Y., Lee, G., Oh, S. M., Yoo, S., & Lee, J. (2021). A Machine Learning Algorithm Predicting Acute Kidney Injury in Intensive Care Unit Patients (NAVOY Acute Kidney Injury): Proof-of-Concept Study. BMC Medical Informatics and Decision Making, 21(1), 368. https://doi.org/10.1186/s12911-021-01740-2

2. Li, Y., Li, M., Wang, Y., Cheng, X., Zhang, Y., & Xue, H. (2023). Predicting in-hospital outcomes of patients with acute kidney injury: a machine learning approach. International Urology and Nephrology, 55, 2055-2066. https://doi.org/10.1007/s11255-023-03646-6

3. Ye, Y., Qian, J., Lu, Y., Ma, Y., Zeng, X., Wen, Y., Zhang, W., Zhu, C., Bai, X., Gao, L., Zhang, Y., Shen, F., & Pan, L. (2023). Prediction of acute kidney injury in patients with liver cirrhosis using machine learning models: evidence from the MIMIC-III and MIMIC-IV. Nature Communications, 14(1), 3800. https://doi.org/10.1038/s41467-023-39474-6

4. Hu, Y., Jiang, X., Ren, H., & Hu, C. (2024). Predictive model of acute kidney injury in critically ill patients with acute pancreatitis: a machine learning approach using the MIMIC-IV database. Renal Failure, 46(1), 2303395. https://doi.org/10.1080/0886022X.2024.2303395

5. Flechet, M., Güiza, F., Schetz, M., Wouters, P., Vanhorebeek, I., Derese, I., Gunst, J., Spriet, I., Casaer, M., Van den Berghe, G., & Meyfroidt, G. (2019). AKIpredictor, an online prognostic calculator for acute kidney injury in adult critically ill patients: development, validation and comparison to serum neutrophil gelatinase-associated

참고문헌

6. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825-2830. https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf

7. Tomašev, N., Glorot, X., Rae, J. W., Zielinski, M., Askham, H., Saraiva, A., Mottram, A., Meyer, C., Ravuri, S., Protsyuk, I., Connell, A., Hughes, C. O., Karthikesalingam, A., Cornebise, J., Montgomery, H., Rees, G., Laing, C., Baker, C. R., Peterson, K., … Mohamed, S. (2019). A clinically applicable approach to continuous prediction of future acute kidney injury. Nature, 572(7767), 116–119. https://doi.org/10.1038/s41586-019-1390-1

8. Mohamadlou, H., Lynn-Palevsky, A., Barton, C., Chettipally, U., Shieh, L., Calvert, J., Saber, N. R., & Das, R. (2018). Prediction of Acute Kidney Injury With a Machine Learning Algorithm Using Electronic Health Record Data. Canadian Journal of Kidney Health and Disease, 5, 2054358118776326. https://doi.org/10.1177/2054358118776326