

「2024 날씨 빅데이터 콘테스트」 수상 후보작 주요 내용

□ 과제 1(수치모델 앙상블을 활용한 강수량 예측)

연번	접수번호	대표자	분석기법	주요 내용
1	240183	현○호	soft voting 앙상블 모델	smote, smoteen 기법 활용 catboost를 활용한 5가지 모델 soft voting
2	240198	최○성	RandomForest 모델	다양한 파생변수 생성, 이진 분류 후 다중 분류 모델 구축
3	240210	정○룡	Bidirection LSTM 모델	단일 확률 변수 등 파생변수 생성
4	240304	윤○윤	stacking 앙상블	단일 확률 파생변수 생성, bagging, lgbm, decision tree 등 5개 모델의 예측결과를 결합하여 메타모델(선형회귀)로 학습
5	240313	강○현	앙상블 모델	lgbm, extratrees, random forest, grandient boosting regressor, huber 모델 앙상블
6	240349	권○연	RandomForest 모델	누적 확률 데이터를 예측 시간별 가중치로 고려하여 재구성
7	240364	김○미	RandomForest 모델	smote 기법 사용
8	240449	이○성	RandomForest 와 엑스트라 트리를 활용한 Soft Voting	절기 등 파생 변수 사용, 저사양 가상환경을 고려하여 모델 구성
9	240480	이○리	RandomForest 모델	장마 변수, v01~v09 변수 통계치 파생 변수 생성
10	240491	전○섭	Logistic 회기 모델	DH 변수 값이 클수록 가중치를 적게 주어 영향력을 줄임, 단일 확률 파생변수 생성

□ 과제2(기상특성에 따른 안개 발생 진단)

연번	접수번호	대표자	분석기법	주요 내용
1	240041	오○욱	XGBoost 모델	다양한 파생변수 생성, 랜덤 언더 샘플링 기법 활용, pca 적용
2	240158	이○규	stacking 앙상블	기상 자료를 이용한 다양한 파생변수 생성, xgboost, catboost, lgbm을 이용한 앙상블
3	240184	이○현	XGBoost 모델	이슬점 온도 등 파생변수 생성, 시정값을 예측 후 시정구간 도출
4	240195	신○균	lightgbm과 xgboost를 이용한 soft voting 모델	pca를 이용한 차원축소, 베이지안 최적화 기법 활용/ 이진 분류 후 다중 분류
5	240247	성○숨	XGBoost 모델	오버 샘플링 기법 사용, 이진 분류 후 다중 분류
6	240253	김○운	catboost 모델	다양한 파생변수 생성
7	240257	박○윤	XGBoost 모델	다양한 파생변수 생성, smote 기법 사용
8	240274	조○우	lgbm 모델	안개 유무 이진분류 후 구간 분류 분석 진행
9	240356	심○연	catboost 모델	기상 자료를 이용한 다양한 파생변수 생성, 언더 샘플링 기법 사용, 지역별로 변수 설정하여 모델 학습
10	240410	성○빈	XGBoost 모델	기상 자료를 이용한 다양한 파생변수 생성

□ 과제3(기상현상과 화재 발생에 대한 상관분석)

연번	접수번호	대표자	분석기법	주요 내용
1	240018	김○재	XGBoost 모델	지역 특성을 기반한 상관분석
2	240137	두○재	RandomForest, XGBoost, Logistic 회기 모델	공간적 자기 상관성을 고려한 파생변수 생성
3	240148	신○원		화재사고 인명피해, 재산피해와 기상현상의 상관성 분석
4	240157	안○환	의사결정나무, RandomForest, k-최근접 이웃, 배깅, 나이브 베이즈를 이용한 소프트 보팅	통계기법 기반 상관분석, 머신러닝 기반 상관분석, 군집기반 상관성 분석
5	240172	류○현	LGBM, CatBoost, Ridge Regression, Random Forest, K-means Clustering	화재 발생 수에 영향을 주는 기상 데이터 분석, 특정 기상 조건, 특정 지역에서 화재 발생 유형과 피해 규모 파악
6	240232	김○윤		기상 요소와 화재 재산 피해 간 상관분석
7	240254	김○선		화재 장소, 원인 등에 따른 기상현상과의 상관성 분석
8	240342	한○원	ANOVA, Extratree, K-means Clustering	강원도 지역 군집 분석을 통한 신 사계절 군집 기반 재산피해와 기상 상관성 분석
9	240444	김○윤		건축물, 임야 유형에 대한 기상현상과 화재 발생 상관성 분석
10	240529	송○호	XGBoost 모델	표준 강수지수, ASOS, AWS, 온실가스 데이터 추가 사용, 기상 데이터와 기후 변화 지표를 활용한 지역별 화재발생 상관분석

□ 과제4(기상에 따른 공동주택 전력수요 예측 개선)

연번	접수번호	대표자	분석기법	주요 내용
1	240106	장○익	XGBoost 모델	아파트 평균 매매 가격, 가구 평균 전력 사용량 데이터 추가 사용, 다양한 파생변수 생성
2	240114	김○희	XGBoost 모델, lgbm 모델	다양한 파생변수 생성
3	240169	안○환	lgbm 모델	pca를 통한 차원 축소 기법
4	240182	신○우	catboost 모델	봄, 겨울과 여름 겨울로 나누어 모델 구축
5	240228	김○지	XGBoost 모델	다양한 파생변수 생성, xgboost 가중 평균 앙상블
6	240248	권○근	XGBoost 모델, lgbm, cat 앙상블 모델	다양한 파생변수 생성
7	240267	이○효	lgbm, XGBoost 앙상블 모델	다양한 파생변수 생성, 계절X지역 군집별 모델 구축
8	240398	정○철	XGBoost 모델, lgbm 모델	불쾌지수 등 파생 변수 생성, xgboost, lgbm 예측 결과 평균을 최종 예측으로 사용하는 앙상블 모델 구축
9	240549	곽○진	XGBoost 모델	이동 평균 등 파생변수 생성, 계절, 지역별로 나눠서 모델 구축
10	240555	최○원	lgbm, catboost 모델	다양한 파생 변수 사용, 시계열적 요소 파악과 잔여 오차 추정에 대한 1차 2차 모델 구성