

Test_0825



테스트 공통 조건

- 클러스터 통일
- 변수 유지(work_time, ...)



테스트 항목 : 변수 추가

▼ 전력소비량 관련 파생변수 추가

- 추가 변수
 - 건물별, 요일별, 시간대별 평균
 - 건물별, 시간대별 평균
 - 건물별, 시간대별 표준편차



테스트 항목 : 변수 선택

▼ feature importance 0인 변수 제외

▼ 사용 변수

- Cluster 0

```
['weekday_hour_mean', 'sy_en', 'date_sin', 'ess_capacity', 'building_number', 'work_time', 'weekday_cos',  
'building_type_Research Institute', 'cooling_area', 'holiday', 'temperature_mean', 'hour_std',  
'building_type_University', 'solar_power_capacity', 'particular', 'total_area', 'discomfort', 'hour_mean',  
'discomfort_mean', 'hour_cos', 'CDH_mean', 'weekday_sin', 'CDH', 'ma_dis_5', 'ma_dis_3', 'pcs_capacity', 'low_da  
y',  
'hour_sin', 'humidity', 'temperature_3', 'sensory_temperature', 'windspeed', 'discomfort_3', 'temperature']
```

- Cluster 1

```
['weekday_hour_mean', 'date_sin', 'hour_mean', 'discomfort', 'work_time', 'weekday_sin', 'total_area', 'CDH_mean',  
'building_number', 'temperature_mean', 'hour_std', 'discomfort_mean', 'cooling_area', 'humidity', 'ma_dis_5',  
'weekday_cos', 'holiday', 'ma_dis_3', 'CDH', 'hour_sin', 'sensory_temperature', 'temperature', 'discomfort_3',  
'hour_cos', 'temperature_3', 'windspeed', 'solar_power_capacity']
```

- Cluster 2

```
['weekday_hour_mean', 'building_number', 'hour_mean', 'date_sin', 'ma_dis_5', 'cooling_area', 'discomfort',  
'hour_sin', 'total_area', 'hour_std', 'CDH', 'hour_cos', 'humidity', 'holiday', 'discomfort_mean', 'CDH_mean',  
'weekday_cos', 'weekday_sin', 'temperature_mean', 'sensory_temperature', 'temperature_3', 'ma_dis_3', 'work_tim  
e',  
'windspeed', 'discomfort_3', 'temperature']
```

- Cluster 3

```
['weekday_hour_mean', 'date_sin', 'holiday', 'hour_std', 'weekday_sin', 'hour_mean', 'total_area', 'discomfort',  
'ma_dis_5', 'low_day', 'CDH', 'building_type_Commercial', 'ma_dis_3', 'cooling_area', 'building_number',  
'sensory_temperature', 'building_type_Hospital', 'temperature_mean', 'CDH_mean', 'weekday_cos', 'discomfort_mea  
n',  
'solar_power_capacity', 'work_time', 'sy_en', 'ess_capacity', 'hour_cos', 'sy_ey', 'hour_sin', 'humidity',  
'windspeed', 'temperature_3', 'discomfort_3', 'temperature']
```

▼ 테스트 : Pycaret - et 지정 모델

- Cluster 0
정규화 수정 전 - SMAPE : 4.08768
정규화 수정 후 - SMAPE : 4.08667
- Cluster 1
정규화 수정 전 - SMAPE : 4.67567
정규화 수정 후 - SMAPE : 4.67014
- Cluster 2
정규화 수정 전 - SMAPE : 2.31264
정규화 수정 후 - SMAPE : 2.29887
- Cluster 3
정규화 수정 전 - SMAPE : 2.93476
정규화 수정 후 - SMAPE : 2.93204

▼ 테스트 : Pycaret - compare model

- Cluster 0 : et
정규화 수정 전 - SMAPE : 4.09256
정규화 수정 후 - SMAPE : 4.09870
- Cluster 1 : et
정규화 수정 전 - SMAPE : 4.69697
정규화 수정 후 - SMAPE : 4.71004
- Cluster 2 : Catboost
정규화 수정 전 - SMAPE : 2.19484
정규화 수정 후 - SMAPE : 2.20908
- Cluster 3 : et
정규화 수정 전 - SMAPE : 2.94628
정규화 수정 후 - SMAPE : 2.93667

[테스트 결과 업데이트 예정]

▼ Shap 결과 전부 사용

▼ 사용 변수

- Cluster 0

```
['weekday_hour_mean', 'date_sin', 'CDH', 'temperature_mean', 'building_number', 'hour_std', 'hour_mean', 'discomfort_mean', 'discomfort', 'holiday', 'CDH_mean', 'ma_dis_3', 'hour_sin', 'total_area', 'cooling_area', 'weekday_sin', 'weekday_cos', 'ma_dis_5', 'solar_power_capacity', 'temperature']
```

- Cluster 1

```
['weekday_hour_mean', 'date_sin', 'discomfort', 'temperature_mean', 'discomfort_mean', 'hour_std', 'CDH_mean', 'hour_mean', 'CDH', 'building_number', 'humidity', 'total_area', 'cooling_area', 'ma_dis_5', 'temperature', 'hour_sin', 'weekday_cos', 'ma_dis_3', 'weekday_sin', 'sensory_temperature']
```

- Cluster 2

```
['weekday_hour_mean', 'hour_mean', 'building_number', 'date_sin', 'ma_dis_5', 'discomfort', 'CDH', 'hour_sin', 'hour_std', 'humidity', 'total_area', 'cooling_area', 'discomfort_mean', 'weekday_cos', 'hour_cos', 'holiday', 'CDH_mean', 'temperat
```

```
'temperature_mean', 'ma_dis_3']
```

- Cluster 3

```
['weekday_hour_mean', 'date_sin', 'holiday', 'hour_mean', 'hour_std', 'CDH', 'sensory_temperature', 'discomfort',  
'hour_sin', 'weekday_sin', 'weekday_sin', 'temperature_mean', 'ma_dis_5', 'total_area', 'building_number',  
'CDH_mean', 'discomfort_mean', 'ma_dis_3', 'cooling_area', 'weekday_cos', 'humidity']
```

▼ 테스트 : Pycaret - et 지정 모델

- Cluster 0

SMAPE : 4.08973

- Cluster 1

SMAPE : 4.66345

- Cluster 2

SMAPE : 2.27170

- Cluster 3

SMAPE : 2.94844

▼ 테스트 : Pycaret - compare model

- Cluster 0 : et

SMAPE : 4.07536

- Cluster 1 : et

SMAPE : 4.65229

- Cluster 2 : CatBoost

SMAPE : 2.12017

- Cluster 3

SMAPE : 2.95855

▼ Shap 결과 내 맘대로 제거

▼ 사용 변수

- Cluster 0

```
['weekday_hour_mean', 'date_sin', 'CDH', 'temperature_mean', 'building_number',  
'hour_std', 'hour_mean', 'discomfort_mean', 'discomfort', 'holiday', 'CDH_mean',  
'ma_dis_3', 'hour_sin', 'total_area', 'cooling_area', 'cooling_area', 'weekday_sin',  
'weekday_cos', 'ma_dis_5', 'solar_power_capacity']
```

- Cluster 1

```
['weekday_hour_mean', 'date_sin', 'discomfort', 'temperature_mean', 'discomfort_mean',  
'hour_std', 'CDH', 'building_number', 'humidity', 'total_area', 'ma_dis_5', 'temperature',  
'hour_sin', 'weekday_cos', 'ma_dis_3']
```

- Cluster 2

```
['weekday_hour_mean', 'hour_mean', 'building_number', 'date_sin', 'ma_dis_5',  
'discomfort', 'CDH', 'hour_sin', 'hour_std', 'humidity', 'total_area']
```

- Cluster 3

```
['weekday_hour_mean', 'date_sin', 'holiday', 'hour_mean', 'hour_std', 'CDH',  
'sensory_temperature', 'discomfort', 'hour_sin', 'weekday_sin', 'temperature_mean',  
'ma_dis_5', 'total_area', 'building_number', 'CDH_mean', 'discomfort_mean']
```

▼ 테스트 : Pycaret - et 지정 모델

- Cluster 0

SMAPE : 4.05002

- Cluster 1

SMAPE : 4.73512

- Cluster 2

SMAPE : 2.43135

- Cluster 3

SMAPE : 2.92323

▼ 테스트 : Pycaret - compare model

- Cluster 0 : et

SMAPE : 4.05499

- Cluster 1 : et

SMAPE : 4.69945

- Cluster 2 : CatBoost

SMAPE : 2.27453

- Cluster 3 : et

SMAPE : 2.92448