

KML Challenge

Secret of 136 Submissions

Table of Contents

1차 과정



Feature Engineering

80회 진행

Model Selection

5회 진행

Ensemble

산술 vs 기하평균
20번 진행

2차 과정



Reference Kernal Code

앞 과정과 적합한 커널 참고
커널 코드를 적절히 변환 - 15번 진행

Model Tuning

Random Search
6회 진행

A/B Test

최적해 찾기
10번 진행

Model Selection

파라미터 튜닝 전

- Gradient Boosting Classifier
- XGBClassifier
- LGBMClassifier

Ensemble

파라미터 튜닝 전

- Voting Classifier : Gradient Boosting Classifier, XGBClassifier, LGBMClassifier
3개의 모델을 앙상블하였다.
- Voting, Gradient, XGB, LGBM 산술/기하 평균 비교
- 제출 결과 Score가 높은 기법 선택

Model Tuning

| | 튜닝 전 | 튜닝 후 |
|---------------|----------------------|---------|
| Public Score | 0.71385 | 0.71836 |
| Private Score | ?? | 0.72426 |
| 튜닝시간 | 모델 3개 튜닝 약 80분 소요 | |

- 랜덤서치로 하이퍼 파라미터 튜닝
- Public Score로 비교 결과 튜닝한 모델이 성능이 좋아짐

A/B Test

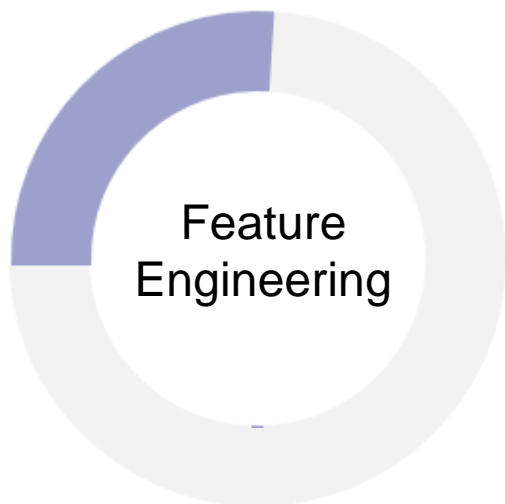
| Pred a | Pred b | Pred c | Pred d | Pred e | Pred f | Pred g | Score |
|--------|--------|--------|--------|--------|--------|--------|---------|
| o | o | o | o | o | o | o | 0.71778 |
| o | o | o | o | o | o | x | 0.71510 |
| o | o | o | o | o | x | o | 0.71348 |
| o | o | o | o | x | o | o | 0.71601 |
| o | o | o | x | o | o | o | 0.71908 |
| o | o | x | o | o | o | o | 0.71938 |
| o | x | o | o | o | o | o | 0.71819 |
| x | o | o | o | o | o | o | 0.71514 |
| o | x | x | x | o | o | o | 0.72069 |

- A/B Test 결과

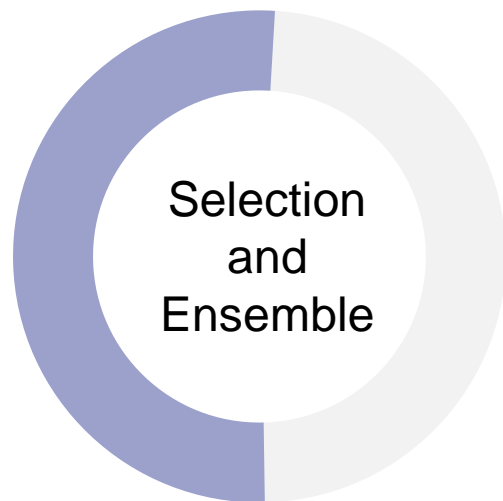
1. 전체 Pred로 기하평균 후 제출
2. Pred를 하나씩 제거하면서 제출
3. Pred 제거시 Score ↑ : 제거
Pred 제거시 Score ↓ : 유지
4. 최종 결과 a, e, f, g 선택

결론

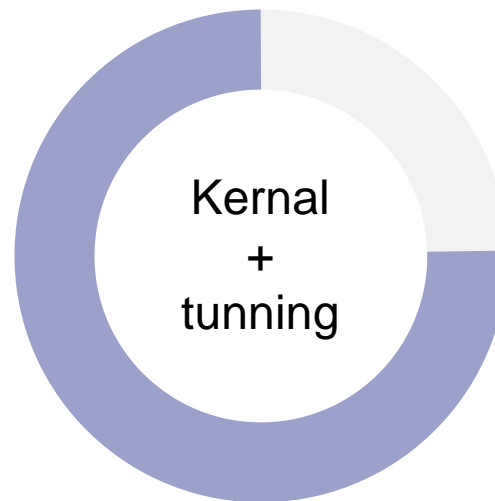
Start : 0.67364



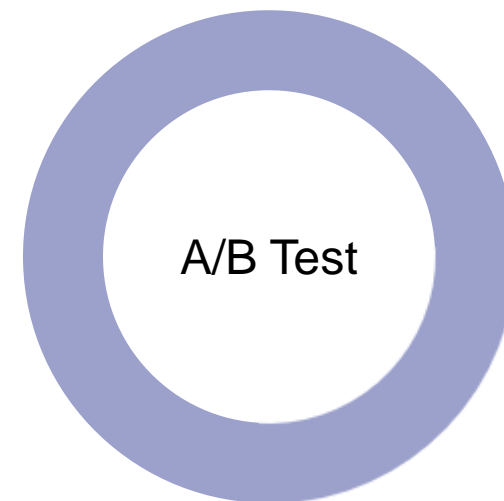
0.035↑



0.002↑



0.007↑



0.003↑