

# LightGBMパラメータチューニング

# Optunaとは？

オープンソースのハイパーパラメータ自動最適化フレームワークOptuna™は、ハイパーパラメータの値に関する試行錯誤を自動化し、優れた性能を発揮するハイパーパラメータの値を自動的に発見します。オープンソースの深層学習フレームワークChainerをはじめ、様々な機械学習ソフトウェアと一緒に使用することが可能です。

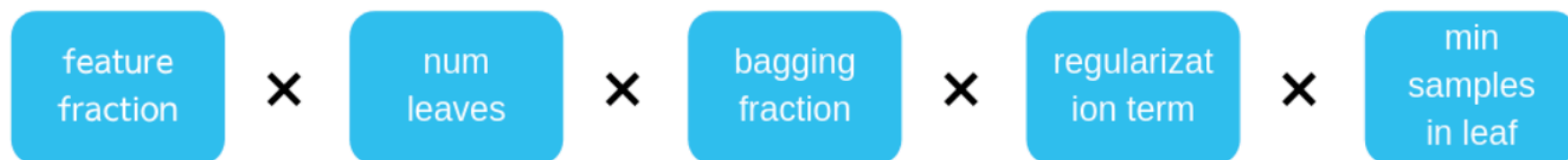
[Preferred Networks, Inc. Optuna公式サイトより引用](#)

## Optunaの強み

- 検索空間と目的関数を別に記述することで可読性・管理性の向上
- 見込みのないパラメータを早期で枝切りするアルゴリズム
- 並列処理化・チューニング過程を保存することで中断再開が可能

## Optuna(旧バージョン)の弱み

- 探索したいハイパーパラメータを組み合わせで探索する為、探索空間が膨大になってしまう（下図）



```
param = {
    'objective': 'binary',
    'metric': 'binary_logloss',
    'lambda_l1': trial.suggest_loguniform('lambda_l1', 1e-8, 1e-4),
    'lambda_l2': trial.suggest_loguniform('lambda_l2', 1e-8, 1e-4),
    'num_leaves': trial.suggest_int('num_leaves', 2, 256),
    'feature_fraction': trial.suggest_uniform('feature_fraction', 0.5, 1.0),
    'bagging_fraction': trial.suggest_uniform('bagging_fraction', 0.5, 1.0),
    'bagging_freq': trial.suggest_int('bagging_freq', 1, 7),
    'min_child_samples': trial.suggest_int('min_child_samples', 10, 100),
}

gbm = lgb.train(param, dtrain)
preds = gbm.predict(test_x)
pred_labels = np rint(preds)
accuracy = sklearn.metrics.accuracy_score(test_y, pred_labels)
return accuracy
```

このパラメータを    この範囲で探索してください



# 実行してみると

```
tune_feature_fraction, val_score: inf: 0%|                                     | 0/7 [00:00<?, ?it/s]

Training until validation scores don't improve for 100 rounds
[50]   valid_0's rmse: 5.48783
[100]   valid_0's rmse: 4.34941
[150]   valid_0's rmse: 4.03509
[200]   valid_0's rmse: 3.95396
[250]   valid_0's rmse: 3.93174
[300]   valid_0's rmse: 3.9458
Early stopping, best iteration is:
[249]   valid_0's rmse: 3.93161

tune_feature_fraction, val_score: 3.931614: 14%|#####7                    | 1/7 [00:00<00:03, 1.54it/s] [I 2020-01-14 20:25:0
9,138] Finished trial#0 resulted in value: 3.9316139524095592. Current best value is 3.9316139524095592 with parameters: {'feature_fractio
n': 0.4}.
tune_feature_fraction, val_score: 3.931614: 14%|#####7                    | 1/7 [00:00<00:03, 1.54it/s]

Training until validation scores don't improve for 100 rounds
[50]   valid_0's rmse: 5.53013
[100]   valid_0's rmse: 4.6255
[150]   valid_0's rmse: 4.34658
[200]   valid_0's rmse: 4.2499
[250]   valid_0's rmse: 4.17927
[300]   valid_0's rmse: 4.13969
[350]   valid_0's rmse: 4.11437
[400]   valid_0's rmse: 4.10375
[450]   valid_0's rmse: 4.10293
[500]   valid_0's rmse: 4.09703
[550]   valid_0's rmse: 4.09988
[600]   valid_0's rmse: 4.10899
Early stopping, best iteration is:
[514]   valid_0's rmse: 4.0946

tune_feature_fraction, val_score: 3.931614: 29%|#####4                    | 2/7 [00:02<00:04, 1.13it/s] [I 2020-01-14 20:25:1
0,568] Finished trial#1 resulted in value: 4.0946018179085915. Current best value is 3.9316139524095592 with parameters: {'feature_fractio
n': 0.4}.
tune_feature_fraction, val_score: 3.931614: 29%|#####4                    | 2/7 [00:02<00:04, 1.13it/s]

Training until validation scores don't improve for 100 rounds
[50]   valid_0's rmse: 4.72123
[100]   valid_0's rmse: 4.03936
[150]   valid_0's rmse: 3.84849
```

処理時間

旧バージョン

Stepwise Tuning

60分

**1/60**

1分

## まとめ

- LightGBMのハイパーパラメータチューニングについて紹介
- ざっくりとした概念について簡単に解説
- 旧バージョンより圧倒的にパラメータ探索を効率よく行える

## 参考サイト

- <https://rin-effort.com/2020/01/14/machine-learning-9/>
- <https://tech.preferred.jp/ja/blog/hyperparameter-tuning-with-optuna-integration-lightgbm-tuner/>