29-Optuna_Hypertuning

October 20, 2024

1 Hypertuning with Optuna

Optuna is an open-source, hyperparameter optimization framework designed for machine learning and deep learning models. It automates the search for optimal hyperparameters, helping data scientists and machine learning engineers tune models more efficiently. Optuna is flexible, scalable, and supports both sequential and parallel optimization, allowing it to adapt to different scales of problems.

Optuna's key strength is its define-by-run approach, where hyperparameter configurations are dynamically constructed during each trial, making it more efficient and flexible than traditional grid or random search methods.

1.0.1 Why is Optuna Important?

Tuning hyperparameters is a critical part of improving machine learning models' performance, but it can be extremely time-consuming, especially for complex models like deep neural networks or ensembles. Optuna helps by:

- **Automating hyperparameter search**: Reducing the manual effort required to find optimal settings.
- **Efficient search**: Optuna uses sophisticated algorithms like Tree-structured Parzen Estimator (TPE) and CMA-ES (Covariance Matrix Adaptation Evolution Strategy) to explore the hyperparameter space more intelligently than grid or random search.
- **Early stopping**: It includes pruning to stop trials that are not promising early, saving time and computational resources.
- **Scalability**: Optuna can be easily scaled to large clusters or cloud environments, making it practical for tuning complex models.

1.0.2 How does Optuna Hypertuning Work?

Optuna performs hyperparameter optimization through an iterative process called trials, where each trial represents a single set of hyperparameters and their corresponding evaluation. Optuna tries to minimize (or maximize) the objective function defined by the user. Here's how it works:

1. Define the Objective Function: The objective function evaluates the model's performance using a given set of hyperparameters. It returns a metric (like accuracy, loss, etc.) to minimize (or maximize).

2. Sampling Hyperparameters: Optuna uses advanced optimization algorithms like TPE or CMA-ES to select hyperparameter values for each trial. It doesn't just randomly choose values but uses past trial results to guide the search toward more promising regions of the hyperparameter space.

3. Run Trials: Optuna evaluates the model with the selected hyperparameters in each trial and records the objective value (e.g., validation accuracy or loss).

4. Pruning: If a trial is not performing well, Optuna can prune (i.e., stop) it early to save resources and move on to more promising hyperparameter combinations.

5. Repeat: Optuna repeats the trial process multiple times (as defined by the user) until it finds the best set of hyperparameters or the predefined number of trials is completed.

1.0.3 How does Optuna Compare to Other Hyperparameter Optimization Methods?

Method	Pros	Cons
Grid Search	Exhaustively searches through predefined values.	Computationally expensive. Limited by pre-defined grid.
Random Search	Explores the search space randomly and is simpler than grid.	Inefficient for large search spaces. May miss optimal values.
Bayesian Search	More efficient by modeling the function being optimized.	Requires more complex setup. Slower for high-dimensional spaces.
Optuna	Define-by-run, efficient search with pruning and parallelism.	More complex to set up than simple random or grid search.

1.0.4 When should you use Optuna?

- **Hyperparameter Optimization for Complex Models**: If you are working with complex machine learning models like Gradient Boosting, Deep Neural Networks, or ensembles, where manual tuning is impractical.
- **Large Hyperparameter Search Space**: When the number of hyperparameters and possible values is large, Optuna's efficiency makes it superior to grid or random search.
- **Need for Resource Efficiency**: Optuna's ability to prune underperforming trials helps save computational time and resources, making it ideal for long-running tasks or expensive-to-evaluate models.

1.0.5 Who uses Optuna?

- **Machine Learning Engineers and Data Scientists**: Optuna is widely adopted by professionals working on machine learning projects where hyperparameter optimization is crucial for achieving high performance.

- **Deep Learning Researchers**: Researchers often use Optuna for neural network hyperparameter tuning, as it efficiently searches through learning rates, optimizers, and architectures.
- **Kaggle Competitors**: Competitors in machine learning competitions use Optuna to gain an edge by finding the best possible hyperparameters.

1.0.6 Key Features of Optuna:

1. Define-by-Run: Hyperparameter space is defined dynamically during the execution of the trials, allowing for flexibility in how you construct and explore the space.

2. Pruning: Automatically stops unpromising trials to save computation time and focus on better-performing hyperparameter combinations.

3. Parallelism: Optuna supports running multiple trials in parallel across different CPU or GPU resources, speeding up the search process.

4. Visualization: Provides built-in tools for visualizing optimization history, parameter importance, and more, making it easier to analyze the optimization process.

1.0.7 Advantages of Optuna:

- **Efficient Optimization**: Optuna uses state-of-the-art algorithms like TPE for efficient hyperparameter search, making it faster than random or grid search.
- **Dynamic Construction**: Unlike grid search, Optuna builds the hyperparameter space dynamically, making it more flexible.
- **Automatic Pruning**: Unpromising trials are stopped early, saving computation time. Parallelization: Optuna can parallelize trials to speed up the optimization process.
- **Visualization Tools**: Optuna includes tools to visualize the optimization process, which helps in understanding how hyperparameters impact model performance.

[34]:
```
import pandas as pd
import seaborn as sns
import optuna
import matplotlib.pyplot as plt

from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.ensemble import RandomForestRegressor, GradientBoostingRegressor
from sklearn.metrics import mean_absolute_error, root_root_mean_squared_error,␣
 ↪r2_score

healthexp = sns.load_dataset('healthexp')
healthexp = pd.get_dummies(healthexp)
healthexp
```

[34]:

	Year	Spending_USD	Life_Expectancy	Country_Canada	Country_France	\
0	1970	252.311	70.6	False	False	
1	1970	192.143	72.2	False	True	
2	1970	123.993	71.9	False	False	

```
3    1970       150.437             72.0          False           False
4    1970       326.961             70.9          False           False
..    …           …                 …              …              …
269  2020      6938.983             81.1          False           False
270  2020      5468.418             82.3          False            True
271  2020      5018.700             80.4          False           False
272  2020      4665.641             84.7          False           False
273  2020     11859.179             77.0          False           False

     Country_Germany  Country_Great Britain  Country_Japan  Country_USA
0               True                  False          False        False
1              False                  False          False        False
2              False                   True          False        False
3              False                  False           True        False
4              False                  False          False         True
..               …                     …              …            …
269             True                  False          False        False
270            False                  False          False        False
271            False                   True          False        False
272            False                  False           True        False
273            False                  False          False         True

[274 rows x 9 columns]
```

[27]:
```
X = healthexp.drop(['Life_Expectancy'], axis=1)
y= healthexp['Life_Expectancy']

X_train, X_test, y_train, y_test = train_test_split(X,y,test_size=0.2,␣
 ↪random_state=54)
rfr = GradientBoostingRegressor(random_state=34)
rfr.fit(X_train, y_train)
```

[27]:
```
GradientBoostingRegressor(random_state=34)
```

[28]:
```
y_pred = rfr.predict(X_test)
y_pred
```

[28]:
```
array([78.39157822, 79.56433878, 77.29122733, 79.10343937, 80.96445194,
       77.29122733, 81.03357969, 81.0478605 , 73.89333767, 77.36302545,
       81.88363916, 77.31394029, 71.47882512, 74.82462518, 82.95075741,
       78.07791025, 71.37073548, 81.34544275, 73.48856096, 80.50574084,
       75.02644369, 80.19365241, 73.22967666, 78.54255466, 77.87168267,
       76.14733657, 73.14057728, 82.16757383, 74.42529019, 76.12931714,
       76.30302906, 81.41008928, 76.11675626, 78.73694658, 78.5494218 ,
       82.64753618, 81.09723669, 76.11675626, 77.89083117, 76.6816684 ,
       79.75257164, 84.16434938, 74.41614983, 78.72096389, 82.16529143,
       84.03429896, 79.14171978, 80.44485511, 81.71684243, 81.2197126 ,
```

```
80.93125484, 79.89895388, 78.76478282, 78.28311191, 79.3784097 ])
```

[29]:
```
mean_absolute_error(y_test, y_pred)
```

[29]:
```
0.2709170644443816
```

[30]:
```
root_root_mean_squared_error(y_test, y_pred)
```

[30]:
```
0.355971309502974
```

[31]:
```
r2_score(y_test, y_pred)
```

[31]:
```
0.9866397423796623
```

[32]:
```
def objective(trial):
    n_estimators = trial.suggest_int('n_estimators', 50, 300)
    learning_rate = trial.suggest_float('learning_rate', 0.01, 0.3)
    max_depth = trial.suggest_int('max_depth', 2, 10)
    subsample = trial.suggest_float('subsample', 0.5, 1.0)

    # Define the model
    model = GradientBoostingRegressor(
        n_estimators=n_estimators,
        learning_rate=learning_rate,
        max_depth=max_depth,
        subsample=subsample,
        random_state=42
    )

    model.fit(X_train, y_train)

    y_pred = model.predict(X_test)
    mse = root_root_mean_squared_error(y_test, y_pred)

    return mse

study = optuna.create_study(direction='minimize')
study.optimize(objective, n_trials=100)

print(f"Best hyperparameters: {study.best_params}")
```

```
[I 2024-10-20 11:55:27,700] A new study created in memory with name: no-
name-96fbbaf2-3f29-44b6-9cc6-00f7be9dc2b3
[I 2024-10-20 11:55:27,950] Trial 0 finished with value: 0.32519704881303185 and
parameters: {'n_estimators': 297, 'learning_rate': 0.267593117461795,
'max_depth': 9, 'subsample': 0.7706360903802824}. Best is trial 0 with value:
0.32519704881303185.
[I 2024-10-20 11:55:28,018] Trial 1 finished with value: 0.32933339451026583 and
```

```
parameters: {'n_estimators': 95, 'learning_rate': 0.2589727113442577,
'max_depth': 5, 'subsample': 0.8863003094336444}. Best is trial 0 with value:
0.32519704881303185.
[I 2024-10-20 11:55:28,338] Trial 2 finished with value: 0.34532580713847005 and
parameters: {'n_estimators': 279, 'learning_rate': 0.21427980221020418,
'max_depth': 9, 'subsample': 0.9493276445660934}. Best is trial 0 with value:
0.32519704881303185.
[I 2024-10-20 11:55:28,597] Trial 3 finished with value: 0.337141751498068 and
parameters: {'n_estimators': 288, 'learning_rate': 0.13182683300401327,
'max_depth': 8, 'subsample': 0.8868435801052132}. Best is trial 0 with value:
0.32519704881303185.
[I 2024-10-20 11:55:28,797] Trial 4 finished with value: 0.3124102730275322 and
parameters: {'n_estimators': 274, 'learning_rate': 0.20260187408469205,
'max_depth': 5, 'subsample': 0.9062239337295626}. Best is trial 4 with value:
0.3124102730275322.
[I 2024-10-20 11:55:28,974] Trial 5 finished with value: 0.33315273771793247 and
parameters: {'n_estimators': 218, 'learning_rate': 0.028634699851232688,
'max_depth': 7, 'subsample': 0.9546279936392901}. Best is trial 4 with value:
0.3124102730275322.
[I 2024-10-20 11:55:29,152] Trial 6 finished with value: 0.3531645162647396 and
parameters: {'n_estimators': 256, 'learning_rate': 0.2844066049483992,
'max_depth': 6, 'subsample': 0.6782931822279332}. Best is trial 4 with value:
0.3124102730275322.
[I 2024-10-20 11:55:29,264] Trial 7 finished with value: 0.3257614028015756 and
parameters: {'n_estimators': 151, 'learning_rate': 0.19935313467618743,
'max_depth': 6, 'subsample': 0.7203121769063907}. Best is trial 4 with value:
0.3124102730275322.
[I 2024-10-20 11:55:29,363] Trial 8 finished with value: 0.5949002296375471 and
parameters: {'n_estimators': 184, 'learning_rate': 0.023014559915282794,
'max_depth': 2, 'subsample': 0.7777546955416367}. Best is trial 4 with value:
0.3124102730275322.
[I 2024-10-20 11:55:29,414] Trial 9 finished with value: 0.44110445545375904 and
parameters: {'n_estimators': 70, 'learning_rate': 0.04678877022641999,
'max_depth': 4, 'subsample': 0.642435911988464}. Best is trial 4 with value:
0.3124102730275322.
[I 2024-10-20 11:55:29,559] Trial 10 finished with value: 0.283432180253704 and
parameters: {'n_estimators': 233, 'learning_rate': 0.1213112051937812,
'max_depth': 3, 'subsample': 0.5246775038673241}. Best is trial 10 with value:
0.283432180253704.
[I 2024-10-20 11:55:29,727] Trial 11 finished with value: 0.2733841950906861 and
parameters: {'n_estimators': 229, 'learning_rate': 0.11446729346660696,
'max_depth': 3, 'subsample': 0.5082849682777049}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:29,882] Trial 12 finished with value: 0.33738549688910224
and parameters: {'n_estimators': 221, 'learning_rate': 0.1142162775589909,
'max_depth': 2, 'subsample': 0.512223578308002}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:30,016] Trial 13 finished with value: 0.2821326667996498 and
```

parameters: {'n_estimators': 225, 'learning_rate': 0.08900454928810904,
'max_depth': 3, 'subsample': 0.5104345068793334}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:30,118] Trial 14 finished with value: 0.2936606123578851 and
parameters: {'n_estimators': 153, 'learning_rate': 0.07506187156406356,
'max_depth': 4, 'subsample': 0.582380490793184}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:30,236] Trial 15 finished with value: 0.2971268863998212 and
parameters: {'n_estimators': 186, 'learning_rate': 0.08185626842879058,
'max_depth': 3, 'subsample': 0.5830011523959723}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:30,386] Trial 16 finished with value: 0.2784443164517285 and
parameters: {'n_estimators': 244, 'learning_rate': 0.16865472566121278,
'max_depth': 3, 'subsample': 0.5847662153259866}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:30,549] Trial 17 finished with value: 0.32945547373396866
and parameters: {'n_estimators': 251, 'learning_rate': 0.16414146586098646,
'max_depth': 4, 'subsample': 0.5904421244601817}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:30,662] Trial 18 finished with value: 0.3243825595212696 and
parameters: {'n_estimators': 195, 'learning_rate': 0.15154369262583414,
'max_depth': 2, 'subsample': 0.6578053194559447}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:30,797] Trial 19 finished with value: 0.352405750294523 and
parameters: {'n_estimators': 124, 'learning_rate': 0.17254249831711935,
'max_depth': 10, 'subsample': 0.5615022984441036}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:30,950] Trial 20 finished with value: 0.36472528222929507
and parameters: {'n_estimators': 206, 'learning_rate': 0.22891639396218252,
'max_depth': 5, 'subsample': 0.8154315863946258}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:31,100] Trial 21 finished with value: 0.3073020016778538 and
parameters: {'n_estimators': 250, 'learning_rate': 0.08462566454055234,
'max_depth': 3, 'subsample': 0.5022795336584275}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:31,240] Trial 22 finished with value: 0.28500727995486796
and parameters: {'n_estimators': 238, 'learning_rate': 0.09996116056277207,
'max_depth': 3, 'subsample': 0.5454805945154205}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:31,359] Trial 23 finished with value: 0.3241893163539905 and
parameters: {'n_estimators': 162, 'learning_rate': 0.06051969627768748,
'max_depth': 4, 'subsample': 0.6320900822024877}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:31,508] Trial 24 finished with value: 0.3340743120889516 and
parameters: {'n_estimators': 269, 'learning_rate': 0.14404392207024466,
'max_depth': 2, 'subsample': 0.6182634714905018}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:31,646] Trial 25 finished with value: 0.28898722310627706

```
and parameters: {'n_estimators': 215, 'learning_rate': 0.18325206586845688,
'max_depth': 3, 'subsample': 0.6968492329500976}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:31,805] Trial 26 finished with value: 0.2859771349934609 and
parameters: {'n_estimators': 233, 'learning_rate': 0.11248707230834971,
'max_depth': 5, 'subsample': 0.5452022752235064}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:31,942] Trial 27 finished with value: 0.31601034658811294
and parameters: {'n_estimators': 199, 'learning_rate': 0.09525980881503929,
'max_depth': 4, 'subsample': 0.6028687123419023}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:32,049] Trial 28 finished with value: 0.3411566621028202 and
parameters: {'n_estimators': 170, 'learning_rate': 0.1377780919990122,
'max_depth': 2, 'subsample': 0.5441643781159113}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:32,266] Trial 29 finished with value: 0.3058034089127453 and
parameters: {'n_estimators': 293, 'learning_rate': 0.06232785721282394,
'max_depth': 7, 'subsample': 0.5007527369849425}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:32,354] Trial 30 finished with value: 0.34296501275487 and
parameters: {'n_estimators': 127, 'learning_rate': 0.239036660636862,
'max_depth': 3, 'subsample': 0.750523188775213}. Best is trial 11 with value:
0.2733841950906861.
[I 2024-10-20 11:55:32,498] Trial 31 finished with value: 0.2687702975135457 and
parameters: {'n_estimators': 233, 'learning_rate': 0.12255598931314905,
'max_depth': 3, 'subsample': 0.5309087953054625}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:32,650] Trial 32 finished with value: 0.2911019907959925 and
parameters: {'n_estimators': 255, 'learning_rate': 0.10341750383000117,
'max_depth': 3, 'subsample': 0.5610610466342697}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:32,813] Trial 33 finished with value: 0.2939611647216501 and
parameters: {'n_estimators': 235, 'learning_rate': 0.16223175960762215,
'max_depth': 4, 'subsample': 0.999010386111054}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:32,964] Trial 34 finished with value: 0.344139250148219 and
parameters: {'n_estimators': 269, 'learning_rate': 0.13409401544100277,
'max_depth': 2, 'subsample': 0.521813772821312}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:33,109] Trial 35 finished with value: 0.3082259462703164 and
parameters: {'n_estimators': 212, 'learning_rate': 0.04311457745659346,
'max_depth': 5, 'subsample': 0.5707021871807052}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:33,287] Trial 36 finished with value: 0.2845700970493313 and
parameters: {'n_estimators': 224, 'learning_rate': 0.1307828538799112,
'max_depth': 3, 'subsample': 0.6103923191274128}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:33,487] Trial 37 finished with value: 0.31574363020947704
```

and parameters: {'n_estimators': 284, 'learning_rate': 0.18707784889203227,
'max_depth': 5, 'subsample': 0.5353396695041222}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:33,724] Trial 38 finished with value: 0.35073747410206124
and parameters: {'n_estimators': 264, 'learning_rate': 0.15184723264528946,
'max_depth': 9, 'subsample': 0.8315364287992428}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:33,877] Trial 39 finished with value: 0.3071835616924356 and
parameters: {'n_estimators': 238, 'learning_rate': 0.12409019896753956,
'max_depth': 4, 'subsample': 0.6574995767812493}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:34,029] Trial 40 finished with value: 0.4677086522347409 and
parameters: {'n_estimators': 194, 'learning_rate': 0.012764854978969356,
'max_depth': 6, 'subsample': 0.7053418368479953}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:34,173] Trial 41 finished with value: 0.28897035646844416
and parameters: {'n_estimators': 243, 'learning_rate': 0.1181888403545856,
'max_depth': 3, 'subsample': 0.5221660145614726}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:34,301] Trial 42 finished with value: 0.32736784804216373
and parameters: {'n_estimators': 222, 'learning_rate': 0.10679544639401044,
'max_depth': 2, 'subsample': 0.5305061638806614}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:34,463] Trial 43 finished with value: 0.29762694215456287
and parameters: {'n_estimators': 279, 'learning_rate': 0.06518549601844656,
'max_depth': 3, 'subsample': 0.5031059382923468}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:34,610] Trial 44 finished with value: 0.26878256364376213
and parameters: {'n_estimators': 228, 'learning_rate': 0.09139112662408969,
'max_depth': 4, 'subsample': 0.5477457735008036}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:34,751] Trial 45 finished with value: 0.29143930744339114
and parameters: {'n_estimators': 205, 'learning_rate': 0.08965053455365322,
'max_depth': 4, 'subsample': 0.5685473731521724}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:34,885] Trial 46 finished with value: 0.39313205257246503
and parameters: {'n_estimators': 225, 'learning_rate': 0.046345326038113074,
'max_depth': 2, 'subsample': 0.5938485089230439}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:34,947] Trial 47 finished with value: 0.3396645631610075 and
parameters: {'n_estimators': 56, 'learning_rate': 0.0731028791459466,
'max_depth': 6, 'subsample': 0.5526624348425062}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:35,071] Trial 48 finished with value: 0.312967918752511 and
parameters: {'n_estimators': 178, 'learning_rate': 0.20356881669983035,
'max_depth': 4, 'subsample': 0.6369500029479468}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:35,231] Trial 49 finished with value: 0.2833206390868708 and

parameters: {'n_estimators': 262, 'learning_rate': 0.14482048416305823,
'max_depth': 3, 'subsample': 0.5772907570360203}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:35,394] Trial 50 finished with value: 0.3254244477042977 and
parameters: {'n_estimators': 244, 'learning_rate': 0.29049585313894283,
'max_depth': 5, 'subsample': 0.5184295143589501}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:35,554] Trial 51 finished with value: 0.28673224445533974
and parameters: {'n_estimators': 262, 'learning_rate': 0.17198550686886205,
'max_depth': 3, 'subsample': 0.5617611851623718}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:35,709] Trial 52 finished with value: 0.2910215237132724 and
parameters: {'n_estimators': 256, 'learning_rate': 0.14777145647614381,
'max_depth': 3, 'subsample': 0.5884989110350857}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:35,841] Trial 53 finished with value: 0.33663834445763846
and parameters: {'n_estimators': 227, 'learning_rate': 0.12351040479543017,
'max_depth': 2, 'subsample': 0.5352066137495815}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:36,044] Trial 54 finished with value: 0.29805083664401705
and parameters: {'n_estimators': 297, 'learning_rate': 0.09354034571522202,
'max_depth': 4, 'subsample': 0.6228478739475506}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:36,176] Trial 55 finished with value: 0.3186550196558229 and
parameters: {'n_estimators': 212, 'learning_rate': 0.16147731103371912,
'max_depth': 3, 'subsample': 0.5820879890308186}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:36,323] Trial 56 finished with value: 0.33128018709998036
and parameters: {'n_estimators': 247, 'learning_rate': 0.14010660706178468,
'max_depth': 2, 'subsample': 0.6544211258426351}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:36,537] Trial 57 finished with value: 0.28831754387250763
and parameters: {'n_estimators': 275, 'learning_rate': 0.07626085785849052,
'max_depth': 4, 'subsample': 0.6799987666663414}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:36,702] Trial 58 finished with value: 0.320019796192551 and
parameters: {'n_estimators': 190, 'learning_rate': 0.10540528737668065,
'max_depth': 8, 'subsample': 0.5130929258940896}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:36,884] Trial 59 finished with value: 0.3431264075744974 and
parameters: {'n_estimators': 260, 'learning_rate': 0.1846341094850861,
'max_depth': 3, 'subsample': 0.6059995367785499}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:37,009] Trial 60 finished with value: 0.36218200880182144
and parameters: {'n_estimators': 206, 'learning_rate': 0.2706652855831479,
'max_depth': 2, 'subsample': 0.5489548662054436}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:37,151] Trial 61 finished with value: 0.27663318682392724

```
and parameters: {'n_estimators': 232, 'learning_rate': 0.11744139207969508,
'max_depth': 3, 'subsample': 0.5258295173126227}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:37,291] Trial 62 finished with value: 0.282327356204379 and
parameters: {'n_estimators': 232, 'learning_rate': 0.11441097886689146,
'max_depth': 3, 'subsample': 0.5802111287236595}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:37,434] Trial 63 finished with value: 0.2846683813869978 and
parameters: {'n_estimators': 230, 'learning_rate': 0.10964627012259157,
'max_depth': 3, 'subsample': 0.5343856984399099}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:37,581] Trial 64 finished with value: 0.29354989451107105
and parameters: {'n_estimators': 216, 'learning_rate': 0.12765983353767738,
'max_depth': 4, 'subsample': 0.5021350240116306}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:37,720] Trial 65 finished with value: 0.35948926495929134
and parameters: {'n_estimators': 240, 'learning_rate': 0.08687325387700343,
'max_depth': 2, 'subsample': 0.5508074120453469}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:37,876] Trial 66 finished with value: 0.2980607734880508 and
parameters: {'n_estimators': 249, 'learning_rate': 0.0971489217202378,
'max_depth': 3, 'subsample': 0.5213821830496556}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:38,027] Trial 67 finished with value: 0.2757023560730843 and
parameters: {'n_estimators': 199, 'learning_rate': 0.11547805321778493,
'max_depth': 4, 'subsample': 0.872277198440613}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:38,166] Trial 68 finished with value: 0.3212843996999328 and
parameters: {'n_estimators': 181, 'learning_rate': 0.08147559495991766,
'max_depth': 5, 'subsample': 0.9404810225597366}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:38,339] Trial 69 finished with value: 0.2998468920089925 and
parameters: {'n_estimators': 199, 'learning_rate': 0.13222337419778468,
'max_depth': 4, 'subsample': 0.8114060461385173}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:38,511] Trial 70 finished with value: 0.2936321850898209 and
parameters: {'n_estimators': 217, 'learning_rate': 0.09959855855692458,
'max_depth': 4, 'subsample': 0.8906174319977933}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:38,663] Trial 71 finished with value: 0.3031109360473692 and
parameters: {'n_estimators': 229, 'learning_rate': 0.11676696430986833,
'max_depth': 3, 'subsample': 0.8394740356203501}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:38,812] Trial 72 finished with value: 0.30506979109723376
and parameters: {'n_estimators': 235, 'learning_rate': 0.11429279951338023,
'max_depth': 3, 'subsample': 0.7812357657877741}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:38,955] Trial 73 finished with value: 0.299335059932847 and
```

```
parameters: {'n_estimators': 204, 'learning_rate': 0.06796442040304482,
'max_depth': 4, 'subsample': 0.5623236135973059}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:39,103] Trial 74 finished with value: 0.27726456031107777
and parameters: {'n_estimators': 221, 'learning_rate': 0.12350693793514392,
'max_depth': 3, 'subsample': 0.5372381389445163}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:39,259] Trial 75 finished with value: 0.3197026616693903 and
parameters: {'n_estimators': 168, 'learning_rate': 0.15526663355775777,
'max_depth': 7, 'subsample': 0.9211213691341871}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:39,393] Trial 76 finished with value: 0.3610721016867259 and
parameters: {'n_estimators': 221, 'learning_rate': 0.05720600204864207,
'max_depth': 2, 'subsample': 0.8570779651846874}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:39,529] Trial 77 finished with value: 0.3017856322375178 and
parameters: {'n_estimators': 212, 'learning_rate': 0.1742764965575541,
'max_depth': 3, 'subsample': 0.5374811319067782}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:39,603] Trial 78 finished with value: 0.28771354036692975
and parameters: {'n_estimators': 89, 'learning_rate': 0.1369162710195497,
'max_depth': 4, 'subsample': 0.5089338457782696}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:39,758] Trial 79 finished with value: 0.36001397984272665
and parameters: {'n_estimators': 251, 'learning_rate': 0.12366063549935136,
'max_depth': 2, 'subsample': 0.7398744705249802}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:39,909] Trial 80 finished with value: 0.28039318485554693
and parameters: {'n_estimators': 241, 'learning_rate': 0.09222678665030536,
'max_depth': 3, 'subsample': 0.5174340187345003}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:40,055] Trial 81 finished with value: 0.2949761769037094 and
parameters: {'n_estimators': 243, 'learning_rate': 0.08205921520156348,
'max_depth': 3, 'subsample': 0.520242399964311}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:40,195] Trial 82 finished with value: 0.28748086601963063
and parameters: {'n_estimators': 221, 'learning_rate': 0.10383294564680376,
'max_depth': 3, 'subsample': 0.5508761620389389}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:40,380] Trial 83 finished with value: 0.28988642449517504
and parameters: {'n_estimators': 238, 'learning_rate': 0.09184546850695231,
'max_depth': 4, 'subsample': 0.5319059506207203}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:40,510] Trial 84 finished with value: 0.330846000770609 and
parameters: {'n_estimators': 197, 'learning_rate': 0.11003247880130926,
'max_depth': 3, 'subsample': 0.9959906722867444}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:40,638] Trial 85 finished with value: 0.2755188597168557 and
```

parameters: {'n_estimators': 209, 'learning_rate': 0.12065534816826647,
'max_depth': 3, 'subsample': 0.5137576808894949}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:40,784] Trial 86 finished with value: 0.3151901181133613 and
parameters: {'n_estimators': 208, 'learning_rate': 0.12034314215245714,
'max_depth': 5, 'subsample': 0.5105430357013842}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:40,904] Trial 87 finished with value: 0.3394515609516515 and
parameters: {'n_estimators': 186, 'learning_rate': 0.13250523083945367,
'max_depth': 2, 'subsample': 0.5689450129169772}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:41,042] Trial 88 finished with value: 0.279441527189532 and
parameters: {'n_estimators': 226, 'learning_rate': 0.15520164978259704,
'max_depth': 3, 'subsample': 0.5388918894107891}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:41,247] Trial 89 finished with value: 0.35259678053966315
and parameters: {'n_estimators': 227, 'learning_rate': 0.15566931509182913,
'max_depth': 10, 'subsample': 0.5940160321261414}. Best is trial 31 with value:
0.2687702975135457.
[I 2024-10-20 11:55:41,392] Trial 90 finished with value: 0.26756036736615935
and parameters: {'n_estimators': 216, 'learning_rate': 0.14322018406580803,
'max_depth': 4, 'subsample': 0.5406516005689751}. Best is trial 90 with value:
0.26756036736615935.
[I 2024-10-20 11:55:41,535] Trial 91 finished with value: 0.2839898731388246 and
parameters: {'n_estimators': 218, 'learning_rate': 0.16538656442887342,
'max_depth': 4, 'subsample': 0.5557724344694196}. Best is trial 90 with value:
0.26756036736615935.
[I 2024-10-20 11:55:41,669] Trial 92 finished with value: 0.27470824373319225
and parameters: {'n_estimators': 202, 'learning_rate': 0.14311660190415712,
'max_depth': 4, 'subsample': 0.5418480086606893}. Best is trial 90 with value:
0.26756036736615935.
[I 2024-10-20 11:55:41,803] Trial 93 finished with value: 0.29370744684684014
and parameters: {'n_estimators': 202, 'learning_rate': 0.14440344299472405,
'max_depth': 4, 'subsample': 0.5284729481912918}. Best is trial 90 with value:
0.26756036736615935.
[I 2024-10-20 11:55:41,950] Trial 94 finished with value: 0.28045739215083354
and parameters: {'n_estimators': 209, 'learning_rate': 0.1275406340731873,
'max_depth': 5, 'subsample': 0.5477745710819973}. Best is trial 90 with value:
0.26756036736615935.
[I 2024-10-20 11:55:42,075] Trial 95 finished with value: 0.29977071364481417
and parameters: {'n_estimators': 191, 'learning_rate': 0.1382706441772703,
'max_depth': 4, 'subsample': 0.5019503850286139}. Best is trial 90 with value:
0.26756036736615935.
[I 2024-10-20 11:55:42,225] Trial 96 finished with value: 0.29074368761970637
and parameters: {'n_estimators': 212, 'learning_rate': 0.14685314197335717,
'max_depth': 4, 'subsample': 0.5280517474257275}. Best is trial 90 with value:
0.26756036736615935.
[I 2024-10-20 11:55:42,409] Trial 97 finished with value: 0.319604857634209 and

```
parameters: {'n_estimators': 255, 'learning_rate': 0.11867606754400264,
'max_depth': 5, 'subsample': 0.5733044545999941}. Best is trial 90 with value:
0.26756036736615935.
[I 2024-10-20 11:55:42,536] Trial 98 finished with value: 0.2857077027347594 and
parameters: {'n_estimators': 174, 'learning_rate': 0.12736296874721426,
'max_depth': 4, 'subsample': 0.5602314712383062}. Best is trial 90 with value:
0.26756036736615935.
[I 2024-10-20 11:55:42,691] Trial 99 finished with value: 0.2878060321930935 and
parameters: {'n_estimators': 234, 'learning_rate': 0.17614425422064806,
'max_depth': 3, 'subsample': 0.5402212509921593}. Best is trial 90 with value:
0.26756036736615935.

Best hyperparameters: {'n_estimators': 216, 'learning_rate':
0.14322018406580803, 'max_depth': 4, 'subsample': 0.5406516005689751}
```

[33]:
```
study.best_params
```

[33]:
```
{'n_estimators': 216,
 'learning_rate': 0.14322018406580803,
 'max_depth': 4,
 'subsample': 0.5406516005689751}
```

[35]:
```
optuna.visualization.plot_optimization_history(study)
```

[36]:
```
optuna.visualization.plot_parallel_coordinate(study)
```

[38]:
```
optuna.visualization.plot_slice(study, params=['n_estimators', 'learning_rate',␣
↪'max_depth', 'subsample'])
```

[40]:
```
optuna.visualization.plot_param_importances(study)
```