flood-prediction

July 14, 2024

[1]:
```
import numpy as np
import pandas as pd

import matplotlib.pyplot as plt
import seaborn as sns

from sklearn.compose import ColumnTransformer
from sklearn.pipeline import Pipeline
from sklearn.model_selection import cross_val_score, train_test_split
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_squared_error, r2_score, mean_absolute_error

from xgboost import XGBRegressor

import optuna
```

```
/home/user/upgrade/.venv/lib/python3.11/site-packages/tqdm/auto.py:21:
TqdmWarning: IProgress not found. Please update jupyter and ipywidgets. See
https://ipywidgets.readthedocs.io/en/stable/user_install.html
  from .autonotebook import tqdm as notebook_tqdm
```

[2]:
```
df = pd.read_csv('flood.csv')
```

[3]:
```
df.head()
```

[3]:

	MonsoonIntensity	TopographyDrainage	RiverManagement	Deforestation	\
0	3	8	6	6	
1	8	4	5	7	
2	3	10	4	1	
3	4	4	2	7	
4	3	7	5	2	

	Urbanization	ClimateChange	DamsQuality	Siltation	AgriculturalPractices	\
0	4	4	6	2	3	
1	7	9	1	5	5	
2	7	5	4	7	4	
3	3	4	1	4	6	
4	5	8	5	2	7	

```
   Encroachments  …  DrainageSystems  CoastalVulnerability  Landslides  \
0              2  …               10                     7           4
1              4  …                9                     2           6
2              9  …                7                     4           4
3              4  …                4                     2           6
4              5  …                7                     6           5

   Watersheds  DeterioratingInfrastructure  PopulationScore  WetlandLoss  \
0           2                            3                4            3
1           2                            1                1            9
2           8                            6                1            8
3           6                            8                8            6
4           3                            3                4            4

   InadequatePlanning  PoliticalFactors  FloodProbability
0                   2                 6             0.450
1                   1                 3             0.475
2                   3                 6             0.515
3                   6                10             0.520
4                   3                 4             0.475

[5 rows x 21 columns]
```

[4]:
```python
def get_df_info(df):
    print("\n\033[1mShape of DataFrame:\033[0m ", df.shape)
    print("\n\033[1mColumns in DataFrame:\033[0m ", df.columns.to_list())
    print("\n\033[1mData types of columns:\033[0m\n", df.dtypes)

    print("\n\033[1mInformation about DataFrame:\033[0m")
    df.info()

    print("\n\033[1mNumber of unique values in each column:\033[0m")
    for col in df.columns:
        print(f"\033[1m{col}\033[0m: {df[col].nunique()}")

    print("\n\033[1mNumber of null values in each column:\033[0m\n", df.
 ↪isnull().sum())

    print("\n\033[1mNumber of duplicate rows:\033[0m ", df.duplicated().sum())

    print("\n\033[1mDescriptive statistics of DataFrame:\033[0m\n", df.
 ↪describe().transpose())

# Call the function
get_df_info(df)
```

Shape of DataFrame: (50000, 21)

Columns in DataFrame: ['MonsoonIntensity', 'TopographyDrainage', 'RiverManagement', 'Deforestation', 'Urbanization', 'ClimateChange', 'DamsQuality', 'Siltation', 'AgriculturalPractices', 'Encroachments', 'IneffectiveDisasterPreparedness', 'DrainageSystems', 'CoastalVulnerability', 'Landslides', 'Watersheds', 'DeterioratingInfrastructure', 'PopulationScore', 'WetlandLoss', 'InadequatePlanning', 'PoliticalFactors', 'FloodProbability']

Data types of columns:

```
 MonsoonIntensity                    int64
TopographyDrainage                   int64
RiverManagement                      int64
Deforestation                        int64
Urbanization                         int64
ClimateChange                        int64
DamsQuality                          int64
Siltation                            int64
AgriculturalPractices                int64
Encroachments                        int64
IneffectiveDisasterPreparedness      int64
DrainageSystems                      int64
CoastalVulnerability                 int64
Landslides                           int64
Watersheds                           int64
DeterioratingInfrastructure          int64
PopulationScore                      int64
WetlandLoss                          int64
InadequatePlanning                   int64
PoliticalFactors                     int64
FloodProbability                   float64
dtype: object
```

Information about DataFrame:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 50000 entries, 0 to 49999
Data columns (total 21 columns):
 #   Column                           Non-Null Count  Dtype
---  ------                           --------------  -----
 0   MonsoonIntensity                 50000 non-null  int64
 1   TopographyDrainage               50000 non-null  int64
 2   RiverManagement                  50000 non-null  int64
 3   Deforestation                    50000 non-null  int64
 4   Urbanization                     50000 non-null  int64
 5   ClimateChange                    50000 non-null  int64
 6   DamsQuality                      50000 non-null  int64
 7   Siltation                        50000 non-null  int64
```

```
 8   AgriculturalPractices            50000 non-null  int64
 9   Encroachments                    50000 non-null  int64
 10  IneffectiveDisasterPreparedness  50000 non-null  int64
 11  DrainageSystems                  50000 non-null  int64
 12  CoastalVulnerability             50000 non-null  int64
 13  Landslides                       50000 non-null  int64
 14  Watersheds                       50000 non-null  int64
 15  DeterioratingInfrastructure      50000 non-null  int64
 16  PopulationScore                  50000 non-null  int64
 17  WetlandLoss                      50000 non-null  int64
 18  InadequatePlanning               50000 non-null  int64
 19  PoliticalFactors                 50000 non-null  int64
 20  FloodProbability                 50000 non-null  float64
dtypes: float64(1), int64(20)
memory usage: 8.0 MB
```

Number of unique values in each column:

MonsoonIntensity: 17
TopographyDrainage: 18
RiverManagement: 17
Deforestation: 18
Urbanization: 18
ClimateChange: 18
DamsQuality: 17
Siltation: 17
AgriculturalPractices: 17
Encroachments: 18
IneffectiveDisasterPreparedness: 17
DrainageSystems: 18
CoastalVulnerability: 18
Landslides: 17
Watersheds: 17
DeterioratingInfrastructure: 18
PopulationScore: 18
WetlandLoss: 19
InadequatePlanning: 17
PoliticalFactors: 17
FloodProbability: 83

Number of null values in each column:

```
 MonsoonIntensity                  0
TopographyDrainage                 0
RiverManagement                    0
Deforestation                      0
Urbanization                       0
ClimateChange                      0
DamsQuality                        0
Siltation                          0
```

```
AgriculturalPractices              0
Encroachments                      0
IneffectiveDisasterPreparedness    0
DrainageSystems                    0
CoastalVulnerability               0
Landslides                         0
Watersheds                         0
DeterioratingInfrastructure        0
PopulationScore                    0
WetlandLoss                        0
InadequatePlanning                 0
PoliticalFactors                   0
FloodProbability                   0
dtype: int64
```

Number of duplicate rows: 0

Descriptive statistics of DataFrame:

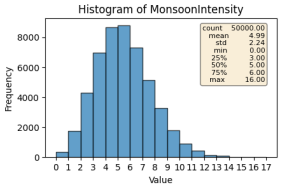
	count	mean	std	min	25%	\
MonsoonIntensity	50000.0	4.99148	2.236834	0.000	3.000	
TopographyDrainage	50000.0	4.98410	2.246488	0.000	3.000	
RiverManagement	50000.0	5.01594	2.231310	0.000	3.000	
Deforestation	50000.0	5.00848	2.222743	0.000	3.000	
Urbanization	50000.0	4.98906	2.243159	0.000	3.000	
ClimateChange	50000.0	4.98834	2.226761	0.000	3.000	
DamsQuality	50000.0	5.01536	2.245000	0.000	3.000	
Siltation	50000.0	4.98860	2.232642	0.000	3.000	
AgriculturalPractices	50000.0	5.00612	2.234588	0.000	3.000	
Encroachments	50000.0	5.00638	2.241633	0.000	3.000	
IneffectiveDisasterPreparedness	50000.0	5.00502	2.226076	0.000	3.000	
DrainageSystems	50000.0	5.00606	2.238107	0.000	3.000	
CoastalVulnerability	50000.0	4.99992	2.247101	0.000	3.000	
Landslides	50000.0	4.98422	2.227741	0.000	3.000	
Watersheds	50000.0	4.97982	2.232190	0.000	3.000	
DeterioratingInfrastructure	50000.0	4.98820	2.231134	0.000	3.000	
PopulationScore	50000.0	4.98498	2.238279	0.000	3.000	
WetlandLoss	50000.0	5.00512	2.231760	0.000	3.000	
InadequatePlanning	50000.0	4.99436	2.230011	0.000	3.000	
PoliticalFactors	50000.0	4.99052	2.246075	0.000	3.000	
FloodProbability	50000.0	0.49966	0.050034	0.285	0.465	

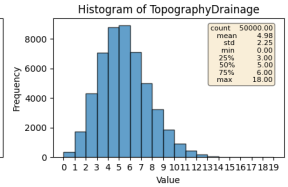
	50%	75%	max
MonsoonIntensity	5.0	6.000	16.000
TopographyDrainage	5.0	6.000	18.000
RiverManagement	5.0	6.000	16.000
Deforestation	5.0	6.000	17.000
Urbanization	5.0	6.000	17.000
ClimateChange	5.0	6.000	17.000

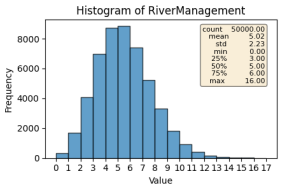
```
DamsQuality                      5.0  6.000  16.000
Siltation                        5.0  6.000  16.000
AgriculturalPractices            5.0  6.000  16.000
Encroachments                    5.0  6.000  18.000
IneffectiveDisasterPreparedness  5.0  6.000  16.000
DrainageSystems                  5.0  6.000  17.000
CoastalVulnerability             5.0  6.000  17.000
Landslides                       5.0  6.000  16.000
Watersheds                       5.0  6.000  16.000
DeterioratingInfrastructure      5.0  6.000  17.000
PopulationScore                  5.0  6.000  19.000
WetlandLoss                      5.0  6.000  22.000
InadequatePlanning               5.0  6.000  16.000
PoliticalFactors                 5.0  6.000  16.000
FloodProbability                 0.5  0.535   0.725
```

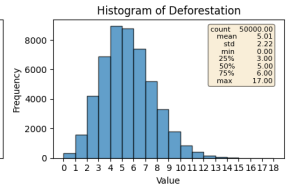
[6]:
```python
# Remove 'FloodProbability' column from the columns list
columns = df.drop('FloodProbability', axis=1).columns
n_cols = 2
n_rows = int(np.ceil(len(columns) / n_cols))

# Create subplots
fig, axes = plt.subplots(nrows=n_rows, ncols=2, figsize=(9, n_rows * 3))
axes = axes.flatten()

# Plot histograms for each column except 'FloodProbability'
for i, column in enumerate(columns):
    ax = axes[i]
    col_data = df[column]
    bins = np.arange(col_data.min(), col_data.max() + 2)
    ax.hist(col_data, bins=bins, alpha=0.7, edgecolor='k')
    ax.set_title(f'Histogram of {column}')
    ax.set_xlabel('Value')
    ax.set_ylabel('Frequency')
    ax.set_xticks(bins)

    # Summary statistics text
    summary_text = col_data.describe().apply(lambda x: f'{x:.2f}').to_string()
    ax.text(0.95, 0.95, summary_text, transform=ax.transAxes, fontsize=8,
            verticalalignment='top', horizontalalignment='right',
            bbox=dict(boxstyle='round', facecolor='wheat', alpha=0.5))

# Remove any unused subplots
for ax in axes[len(columns):]:
    fig.delaxes(ax)

plt.tight_layout()
```

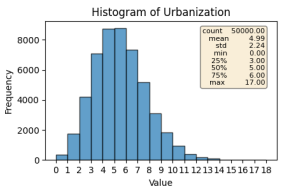
```
plt.show()
```

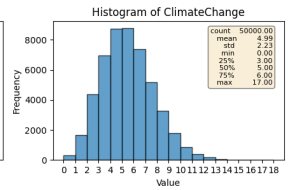

Histogram of MonsoonIntensity
count 50000.00
mean 4.99
std 2.24
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 16.00
Frequency
Value

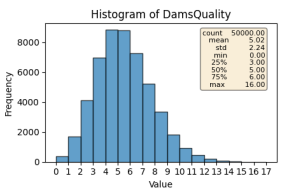

Histogram of TopographyDrainage
count 50000.00
mean 4.98
std 2.25
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 18.00
Frequency
Value

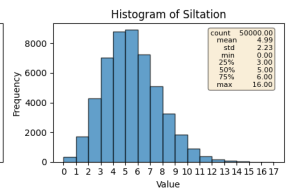

Histogram of RiverManagement
count 50000.00
mean 5.02
std 2.23
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 16.00
Frequency
Value

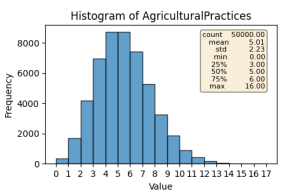

Histogram of Deforestation
count 50000.00
mean 5.01
std 2.22
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 17.00
Frequency
Value

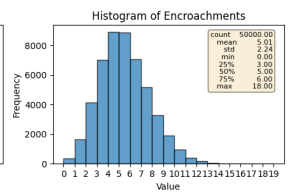

Histogram of Urbanization
count 50000.00
mean 4.99
std 2.24
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 17.00
Frequency
Value

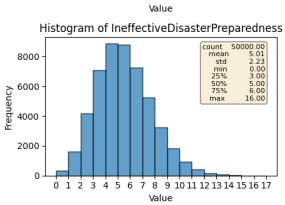

Histogram of ClimateChange
count 50000.00
mean 4.99
std 2.23
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 17.00
Frequency
Value

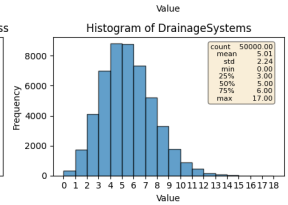

Histogram of DamsQuality
count 50000.00
mean 5.02
std 2.24
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 16.00
Frequency
Value

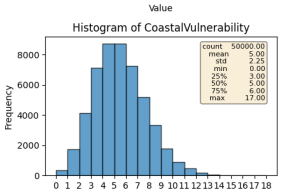

Histogram of Siltation
count 50000.00
mean 4.99
std 2.23
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 16.00
Frequency
Value

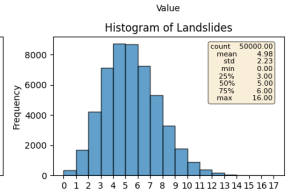

Histogram of AgriculturalPractices
count 50000.00
mean 5.01
std 2.23
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 16.00
Frequency
Value

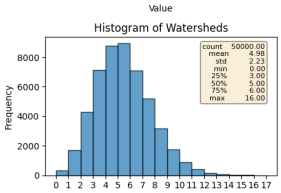

Histogram of Encroachments
count 50000.00
mean 5.01
std 2.24
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 18.00
Frequency
Value

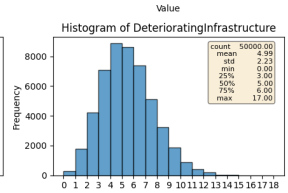

Histogram of IneffectiveDisasterPreparedness
count 50000.00
mean 5.01
std 2.23
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 16.00
Frequency
Value

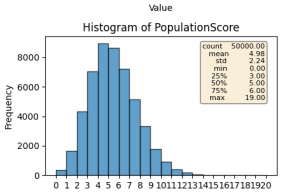

Histogram of DrainageSystems
count 50000.00
mean 5.01
std 2.24
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 17.00
Frequency
Value

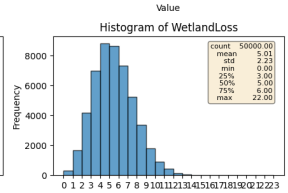

Histogram of CoastalVulnerability
count 50000.00
mean 5.00
std 2.25
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 17.00
Frequency
Value

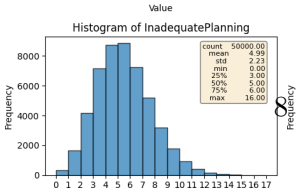

Histogram of Landslides
count 50000.00
mean 4.98
std 2.23
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 16.00
Frequency
Value

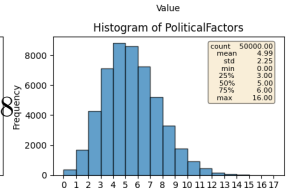

Histogram of Watersheds
count 50000.00
mean 4.98
std 2.23
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 16.00
Frequency
Value


Histogram of DeterioratingInfrastructure
count 50000.00
mean 4.99
std 2.23
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 17.00
Frequency
Value

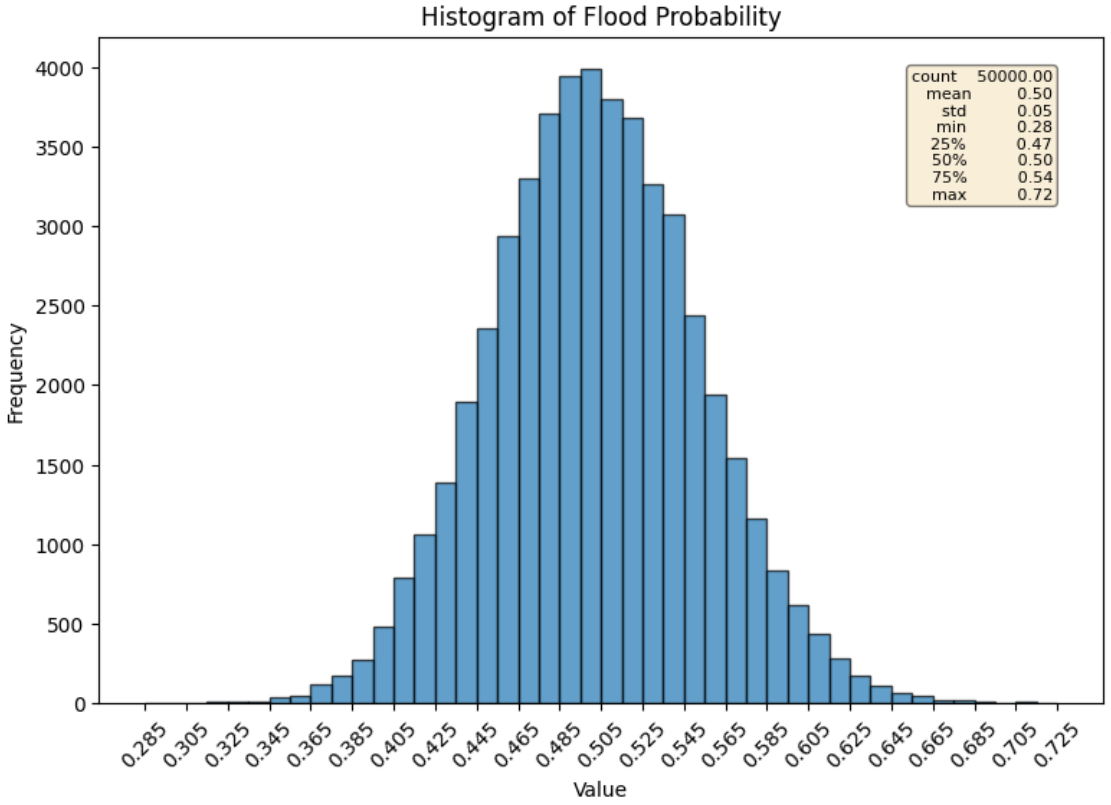

Histogram of PopulationScore
count 50000.00
mean 4.98
std 2.24
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 19.00
Frequency
Value


Histogram of WetlandLoss
count 50000.00
mean 5.01
std 2.23
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 22.00
Frequency
Value

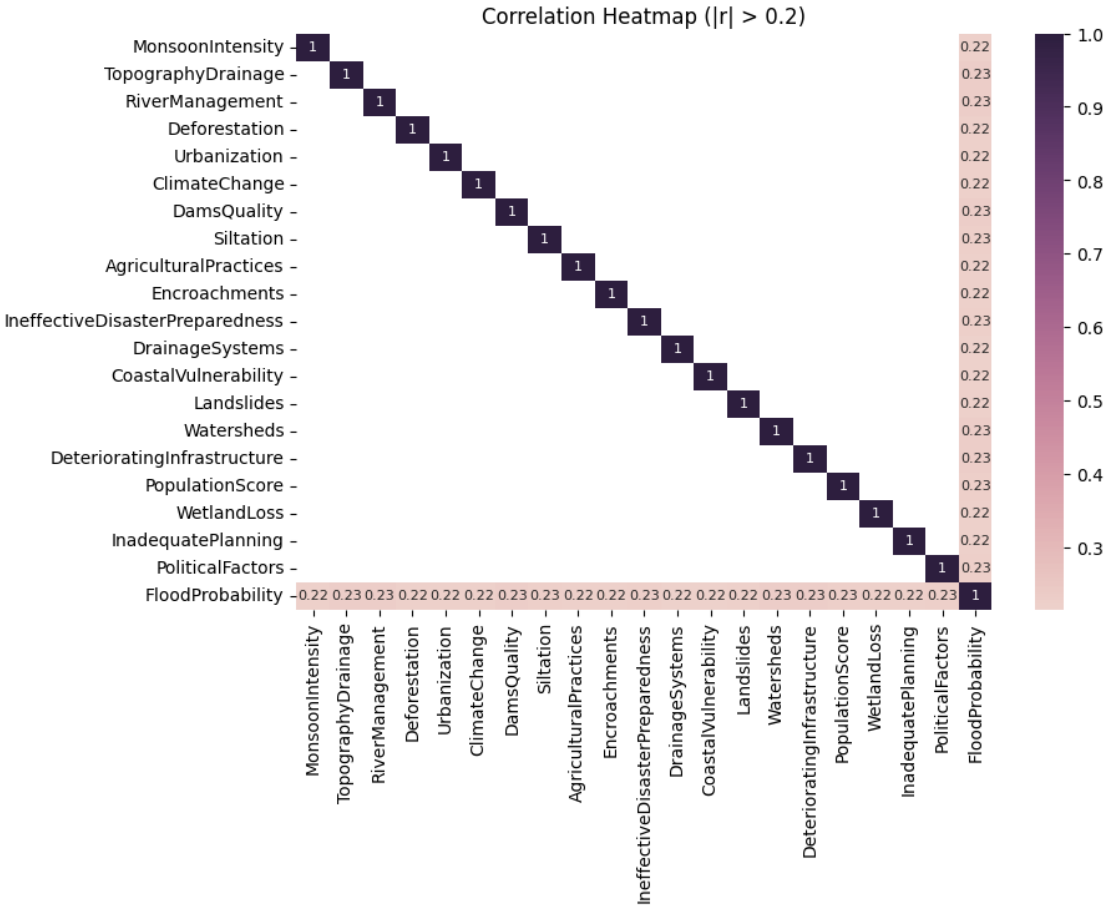

Histogram of InadequatePlanning
count 50000.00
mean 4.99
std 2.23
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 16.00
Frequency
Value


Histogram of PoliticalFactors
count 50000.00
mean 4.99
std 2.25
min 0.00
25% 3.00
50% 5.00
75% 6.00
max 16.00
Frequency
Value

[7]:
```
fig, ax = plt.subplots(figsize=(9, 6))
bins = np.linspace(df['FloodProbability'].min(), df['FloodProbability'].max(),␣
 ↪45) # min, max, bar count

# Plot FloodProbability
ax.hist(df['FloodProbability'], bins=bins, alpha=0.7, edgecolor='k')
ax.set_title(f'Histogram of Flood Probability')
ax.set_xlabel('Value')
ax.set_ylabel('Frequency')
ax.set_xticks(bins[::2])
ax.tick_params(axis='x', rotation=45)

# Summary statistics text
summary_text = df['FloodProbability'].describe().apply(lambda x: f'{x:.2f}').
 ↪to_string()
ax.text(0.95, 0.95, summary_text, transform=ax.transAxes, fontsize=8,
        verticalalignment='top', horizontalalignment='right',
        bbox=dict(boxstyle='round', facecolor='wheat', alpha=0.5))

plt.show()
```



[8]:
```
fig, ax = plt.subplots(figsize=(9, 6))

# Plot Heatmap
corr_matrix = df.corr()
mask = np.abs(corr_matrix) > 0.2
sns.heatmap(corr_matrix, annot=True, cmap=sns.cubehelix_palette(as_cmap=True),␣
 ↪mask=~mask, annot_kws={"size": 8})

plt.title('Correlation Heatmap (|r| > 0.2)')
plt.show()
```



[9]:
```
def preprocess(df):
    df = df.copy()

    X = df.drop('FloodProbability', axis=1)
    y = df['FloodProbability']
```

```
    return X, y
```

[10]:
```
X, y = preprocess(df)
```

[11]:
```
def build_model(params=None):
    numerical_transformer = Pipeline(steps=[
        ('MinMax Scaler', MinMaxScaler())
    ])

    preprocessor = ColumnTransformer(transformers=[
        ('Numerical Transformer', numerical_transformer, X.select_dtypes(np.
 ↪number).columns)
    ], remainder='passthrough')

    if params is None:
        xgb_model = XGBRegressor(random_state=42)
    else:
        xgb_model = XGBRegressor(random_state=42, **params)

    model = Pipeline(steps=[
        ('Preprocess', preprocessor),
        ('XGBModel', xgb_model)
    ])

    return model
```

[12]:
```
model = build_model()

cv_scores = cross_val_score(model, X, y, cv=5,␣
 ↪scoring='neg_root_mean_squared_error')
cv_rmse = -cv_scores.mean()
print(f"Cross-validation RMSE: {cv_rmse}")
```

```
Cross-validation RMSE: 0.013377775099253853
```

[13]:
```
def print_metrics(target, predictions):
    rmse = np.sqrt(mean_squared_error(target, predictions))
    r2 = r2_score(target, predictions)
    mae = mean_absolute_error(target, predictions)
    print(f'Root Mean Squarred Error: {rmse}\nR-squared (r^2): {r2}\nMean␣
 ↪Absolute Error: {mae}')
```

[14]:
```
X, y = preprocess(df)

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.25,␣
 ↪shuffle=True, random_state=42)
model.fit(X_train, y_train)
```

[14]:
```
Pipeline(steps=[('Preprocess',
                 ColumnTransformer(remainder='passthrough',
                                   transformers=[('Numerical Transformer',
                                                  Pipeline(steps=[('MinMax '
                                                                   'Scaler',
MinMaxScaler())]),
                                                  Index(['MonsoonIntensity',
'TopographyDrainage', 'RiverManagement',
       'Deforestation', 'Urbanization', 'ClimateChange', 'DamsQuality',
       'Siltation', 'AgriculturalPractices', 'Encroachments',
       'Ineffecti…
                                 feature_types=None, gamma=None, grow_policy=None,
                                 importance_type=None,
                                 interaction_constraints=None, learning_rate=None,
                                 max_bin=None, max_cat_threshold=None,
                                 max_cat_to_onehot=None, max_delta_step=None,
                                 max_depth=None, max_leaves=None,
                                 min_child_weight=None, missing=nan,
                                 monotone_constraints=None, multi_strategy=None,
                                 n_estimators=None, n_jobs=None,
                                 num_parallel_tree=None, random_state=42, …))])
```

[15]:
```
predictions = model.predict(X_test)
print_metrics(y_test, predictions)
```

```
Root Mean Squarred Error: 0.013584748491394204
R-squared (r^2): 0.9261235596294177
Mean Absolute Error: 0.010802139900875092
```

[16]:
```
X, y = preprocess(df)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.1,␣
 ↪shuffle=True, random_state=42)
X_train, X_val, y_train, y_val = train_test_split(X_train, y_train, test_size=0.
 ↪17, shuffle=True, random_state=42)
```

[17]:
```
def objective(trial):
    params = {
        'max_depth': trial.suggest_int('max_depth', 1, 20),
        'learning_rate': trial.suggest_float('learning_rate', 1e-3, 1.0,␣
 ↪log=True),
        'n_estimators': trial.suggest_int('n_estimators', 100, 2000),
        'min_child_weight': trial.suggest_int('min_child_weight', 1, 10),
        'subsample': trial.suggest_float('subsample', 0.5, 1.0),
        'colsample_bytree': trial.suggest_float('colsample_bytree', 0.5, 1.0),
        'gamma': trial.suggest_float('gamma', 1e-8, 1.0, log=True),
        'reg_alpha': trial.suggest_float('reg_alpha', 1e-8, 1.0, log=True),
        'reg_lambda': trial.suggest_float('reg_lambda', 1e-8, 1.0, log=True),
```

```
    }

    model = build_model(params)
    model.fit(X_train, y_train)

    preds = model.predict(X_val)
    rmse = np.sqrt(mean_squared_error(y_val, preds))
    return rmse
```

[18]:
```
study = optuna.create_study(direction='minimize')
study.optimize(objective, n_trials=100)
```

```
[I 2024-07-14 16:19:06,019] A new study created in memory with name: no-
name-0d98f17d-37a3-488a-8d92-5372c128b427
[I 2024-07-14 16:19:07,905] Trial 0 finished with value: 0.006962904284530168
and parameters: {'max_depth': 5, 'learning_rate': 0.06028437494262665,
'n_estimators': 522, 'min_child_weight': 2, 'subsample': 0.7979492234598261,
'colsample_bytree': 0.7552271095251881, 'gamma': 1.1621644413945642e-06,
'reg_alpha': 4.481078862347643e-08, 'reg_lambda': 0.031519963018996615}. Best is
trial 0 with value: 0.006962904284530168.
[I 2024-07-14 16:19:11,816] Trial 1 finished with value: 0.015516160376300958
and parameters: {'max_depth': 7, 'learning_rate': 0.40181707807843653,
'n_estimators': 748, 'min_child_weight': 4, 'subsample': 0.6087430151648603,
'colsample_bytree': 0.5325806883396416, 'gamma': 5.7515649639032626e-06,
'reg_alpha': 0.0001716123938678449, 'reg_lambda': 0.0035080113007666842}. Best
is trial 0 with value: 0.006962904284530168.
[I 2024-07-14 16:19:18,498] Trial 2 finished with value: 0.04011610245378843 and
parameters: {'max_depth': 20, 'learning_rate': 0.004720979709517702,
'n_estimators': 110, 'min_child_weight': 4, 'subsample': 0.5700385529123448,
'colsample_bytree': 0.7350882613591403, 'gamma': 2.2580704171175567e-05,
'reg_alpha': 3.433488016360125e-06, 'reg_lambda': 4.684950738554976e-06}. Best
is trial 0 with value: 0.006962904284530168.
[I 2024-07-14 16:19:21,738] Trial 3 finished with value: 0.011677748863677851
and parameters: {'max_depth': 15, 'learning_rate': 0.03659074767472406,
'n_estimators': 1058, 'min_child_weight': 6, 'subsample': 0.7448880384860523,
'colsample_bytree': 0.6434877628792854, 'gamma': 0.011729454626835927,
'reg_alpha': 0.0004232713842526916, 'reg_lambda': 2.6871407313949708e-08}. Best
is trial 0 with value: 0.006962904284530168.
[I 2024-07-14 16:19:23,960] Trial 4 finished with value: 0.028244006416553495
and parameters: {'max_depth': 19, 'learning_rate': 0.9388808401293789,
'n_estimators': 1691, 'min_child_weight': 9, 'subsample': 0.5201134935659271,
'colsample_bytree': 0.9482709067805515, 'gamma': 0.09631380036667385,
'reg_alpha': 0.055726952989291555, 'reg_lambda': 0.540047985205865}. Best is
trial 0 with value: 0.006962904284530168.
[I 2024-07-14 16:19:49,438] Trial 5 finished with value: 0.015298994068209946
and parameters: {'max_depth': 13, 'learning_rate': 0.005998612881049525,
'n_estimators': 1887, 'min_child_weight': 10, 'subsample': 0.9476454309951068,
```

'colsample_bytree': 0.9455831842145128, 'gamma': 3.933366468897672e-05,
'reg_alpha': 0.07060207027724603, 'reg_lambda': 0.00013439989731185497}. Best is
trial 0 with value: 0.006962904284530168.
[I 2024-07-14 16:19:57,846] Trial 6 finished with value: 0.01114405076660883 and
parameters: {'max_depth': 13, 'learning_rate': 0.1033321134256728,
'n_estimators': 1282, 'min_child_weight': 10, 'subsample': 0.6126742598888748,
'colsample_bytree': 0.7972808671147533, 'gamma': 1.3770075608796208e-07,
'reg_alpha': 0.22330025388968547, 'reg_lambda': 2.384661051559081e-07}. Best is
trial 0 with value: 0.006962904284530168.
[I 2024-07-14 16:19:58,662] Trial 7 finished with value: 0.012531053220697301
and parameters: {'max_depth': 4, 'learning_rate': 0.519023412541092,
'n_estimators': 460, 'min_child_weight': 9, 'subsample': 0.7457494762345627,
'colsample_bytree': 0.6361617828219766, 'gamma': 0.0016378156656278809,
'reg_alpha': 2.3763515680459578e-06, 'reg_lambda': 6.92975534816569e-06}. Best
is trial 0 with value: 0.006962904284530168.
[I 2024-07-14 16:20:00,916] Trial 8 finished with value: 0.014623901068223156
and parameters: {'max_depth': 15, 'learning_rate': 0.12957640133977333,
'n_estimators': 1488, 'min_child_weight': 4, 'subsample': 0.6648486296786025,
'colsample_bytree': 0.5215173614599664, 'gamma': 0.06811915116541938,
'reg_alpha': 0.00015272215023553776, 'reg_lambda': 0.03847110482620962}. Best is
trial 0 with value: 0.006962904284530168.
[I 2024-07-14 16:20:04,743] Trial 9 finished with value: 0.0375706976647647 and
parameters: {'max_depth': 4, 'learning_rate': 0.0021779205604600628,
'n_estimators': 1091, 'min_child_weight': 2, 'subsample': 0.7174474599038245,
'colsample_bytree': 0.722589336466527, 'gamma': 0.0003258624094002179,
'reg_alpha': 0.0005902990079933001, 'reg_lambda': 0.2682509214040583}. Best is
trial 0 with value: 0.006962904284530168.
[I 2024-07-14 16:20:06,080] Trial 10 finished with value: 0.03650692598940137
and parameters: {'max_depth': 1, 'learning_rate': 0.016832874822130795,
'n_estimators': 648, 'min_child_weight': 1, 'subsample': 0.9024515149415717,
'colsample_bytree': 0.8213875331951956, 'gamma': 2.1085187690900257e-08,
'reg_alpha': 1.8984646370286558e-08, 'reg_lambda': 0.010028199984134587}. Best
is trial 0 with value: 0.006962904284530168.
[I 2024-07-14 16:20:12,084] Trial 11 finished with value: 0.009810255975648836
and parameters: {'max_depth': 9, 'learning_rate': 0.07874294161880915,
'n_estimators': 1363, 'min_child_weight': 7, 'subsample': 0.818127512565686,
'colsample_bytree': 0.8421008114320174, 'gamma': 1.2203103455730393e-07,
'reg_alpha': 0.7312214734015221, 'reg_lambda': 1.2091295712793952e-08}. Best is
trial 0 with value: 0.006962904284530168.
[I 2024-07-14 16:20:13,911] Trial 12 finished with value: 0.012846913551697372
and parameters: {'max_depth': 9, 'learning_rate': 0.06642839933543085,
'n_estimators': 205, 'min_child_weight': 7, 'subsample': 0.8644232759212741,
'colsample_bytree': 0.8706869724147972, 'gamma': 7.976806156107367e-07,
'reg_alpha': 1.612543914538176e-08, 'reg_lambda': 1.171852104475555e-08}. Best
is trial 0 with value: 0.006962904284530168.
[I 2024-07-14 16:20:18,389] Trial 13 finished with value: 0.011552843772533152
and parameters: {'max_depth': 7, 'learning_rate': 0.01499023310692964,
'n_estimators': 795, 'min_child_weight': 7, 'subsample': 0.8405440794780821,

'colsample_bytree': 0.8686448985939085, 'gamma': 7.416856551388588e-07,
'reg_alpha': 1.0268545341796878e-06, 'reg_lambda': 0.011437702718790898}. Best
is trial 0 with value: 0.006962904284530168.
[I 2024-07-14 16:20:26,122] Trial 14 finished with value: 0.013558478896626678
and parameters: {'max_depth': 10, 'learning_rate': 0.10962038784823469,
'n_estimators': 1395, 'min_child_weight': 2, 'subsample': 0.8084145085119266,
'colsample_bytree': 0.6704942218609573, 'gamma': 2.422338330387377e-08,
'reg_alpha': 0.007695608236812884, 'reg_lambda': 1.5147893824088722e-05}. Best
is trial 0 with value: 0.006962904284530168.
[I 2024-07-14 16:20:27,335] Trial 15 finished with value: 0.00777022628331349
and parameters: {'max_depth': 4, 'learning_rate': 0.2446420946485391,
'n_estimators': 415, 'min_child_weight': 5, 'subsample': 0.989527277719954,
'colsample_bytree': 0.7987057303795219, 'gamma': 1.4469458926414334e-06,
'reg_alpha': 1.4659251418652142e-07, 'reg_lambda': 0.0003195894163158505}. Best
is trial 0 with value: 0.006962904284530168.
[I 2024-07-14 16:20:27,996] Trial 16 finished with value: 0.006459144696692934
and parameters: {'max_depth': 1, 'learning_rate': 0.2690598238660217,
'n_estimators': 344, 'min_child_weight': 3, 'subsample': 0.9937742046840934,
'colsample_bytree': 0.7641661298533893, 'gamma': 3.4758510996853956e-06,
'reg_alpha': 2.3997247467542267e-07, 'reg_lambda': 0.0005924836022452638}. Best
is trial 16 with value: 0.006459144696692934.
[I 2024-07-14 16:20:28,633] Trial 17 finished with value: 0.03580020156401153
and parameters: {'max_depth': 1, 'learning_rate': 0.03205158467396898,
'n_estimators': 367, 'min_child_weight': 2, 'subsample': 0.9982973027059051,
'colsample_bytree': 0.587559835844961, 'gamma': 0.00040856821604117507,
'reg_alpha': 1.5037970662410266e-07, 'reg_lambda': 0.049711624650420935}. Best
is trial 16 with value: 0.006459144696692934.
[I 2024-07-14 16:20:30,014] Trial 18 finished with value: 0.005814065861923938
and parameters: {'max_depth': 3, 'learning_rate': 0.23752882152516377,
'n_estimators': 609, 'min_child_weight': 3, 'subsample': 0.9152975596880388,
'colsample_bytree': 0.7523820127728066, 'gamma': 4.972406516221786e-06,
'reg_alpha': 2.0120074231191297e-05, 'reg_lambda': 0.00302200426029384}. Best is
trial 18 with value: 0.005814065861923938.
[I 2024-07-14 16:20:31,388] Trial 19 finished with value: 0.0012202559034307063
and parameters: {'max_depth': 1, 'learning_rate': 0.23973497784017458,
'n_estimators': 910, 'min_child_weight': 3, 'subsample': 0.9178048867402074,
'colsample_bytree': 0.6931642325828917, 'gamma': 9.871435905910376e-06,
'reg_alpha': 1.5710584543901032e-05, 'reg_lambda': 3.3267644056325204e-05}. Best
is trial 19 with value: 0.0012202559034307063.
[I 2024-07-14 16:20:33,858] Trial 20 finished with value: 0.006225366152830098
and parameters: {'max_depth': 3, 'learning_rate': 0.9095583433873078,
'n_estimators': 913, 'min_child_weight': 1, 'subsample': 0.9024252309103982,
'colsample_bytree': 0.6716352240585288, 'gamma': 1.8489590086954664e-05,
'reg_alpha': 2.048669943125516e-05, 'reg_lambda': 2.3294104902600154e-05}. Best
is trial 19 with value: 0.0012202559034307063.
[I 2024-07-14 16:20:35,832] Trial 21 finished with value: 0.006954041489876625
and parameters: {'max_depth': 3, 'learning_rate': 0.9920165125443234,
'n_estimators': 876, 'min_child_weight': 1, 'subsample': 0.89921421045116,

'colsample_bytree': 0.6952655719755207, 'gamma': 1.627115066705433e-05,
'reg_alpha': 1.3924897453871552e-05, 'reg_lambda': 8.94116859344743e-07}. Best
is trial 19 with value: 0.0012202559034307063.
[I 2024-07-14 16:20:37,575] Trial 22 finished with value: 0.010344615651189092
and parameters: {'max_depth': 6, 'learning_rate': 0.22210700318788595,
'n_estimators': 931, 'min_child_weight': 3, 'subsample': 0.9350192786769898,
'colsample_bytree': 0.5888620551672779, 'gamma': 0.0001540410363597707,
'reg_alpha': 2.680123250970087e-05, 'reg_lambda': 3.251924706238294e-05}. Best
is trial 19 with value: 0.0012202559034307063.
[I 2024-07-14 16:20:39,171] Trial 23 finished with value: 0.007682383424310171
and parameters: {'max_depth': 2, 'learning_rate': 0.5177649434086576,
'n_estimators': 1162, 'min_child_weight': 3, 'subsample': 0.8725560124622684,
'colsample_bytree': 0.6851414007139183, 'gamma': 0.0016636355972124291,
'reg_alpha': 2.0064029313337707e-05, 'reg_lambda': 5.119745989817946e-05}. Best
is trial 19 with value: 0.0012202559034307063.
[I 2024-07-14 16:20:40,614] Trial 24 finished with value: 0.006040792263667931
and parameters: {'max_depth': 3, 'learning_rate': 0.18036573188885668,
'n_estimators': 639, 'min_child_weight': 1, 'subsample': 0.9414829020246309,
'colsample_bytree': 0.5806789230090709, 'gamma': 5.75713200158977e-05,
'reg_alpha': 0.0013273613133263943, 'reg_lambda': 1.4466327389320255e-06}. Best
is trial 19 with value: 0.0012202559034307063.
[I 2024-07-14 16:20:41,908] Trial 25 finished with value: 0.009818106939648 43
and parameters: {'max_depth': 6, 'learning_rate': 0.14888767825209812,
'n_estimators': 690, 'min_child_weight': 5, 'subsample': 0.945535291469889,
'colsample_bytree': 0.5734674154385787, 'gamma': 0.0014872460671289261,
'reg_alpha': 0.0037320457314433263, 'reg_lambda': 1.2187192781671242e-06}. Best
is trial 19 with value: 0.0012202559034307063.
[I 2024-07-14 16:20:42,788] Trial 26 finished with value: 0.025649094154880 08
and parameters: {'max_depth': 2, 'learning_rate': 0.1838803754847491,
'n_estimators': 552, 'min_child_weight': 3, 'subsample': 0.9453496973438377,
'colsample_bytree': 0.6126464913316257, 'gamma': 0.5794220778718138,
'reg_alpha': 0.0020246044880796196, 'reg_lambda': 1.2839671764613702e-07}. Best
is trial 19 with value: 0.0012202559034307063.
[I 2024-07-14 16:20:44,541] Trial 27 finished with value: 0.016944444296129204
and parameters: {'max_depth': 8, 'learning_rate': 0.38239602143646584,
'n_estimators': 615, 'min_child_weight': 1, 'subsample': 0.7891634003442166,
'colsample_bytree': 0.7168568687625058, 'gamma': 6.917886403584963e-05,
'reg_alpha': 0.0011197492363095 43, 'reg_lambda': 0.0034169149799829768}. Best is
trial 19 with value: 0.0012202559034307063.
[I 2024-07-14 16:20:45,543] Trial 28 finished with value: 0.010059491829998 95
and parameters: {'max_depth': 5, 'learning_rate': 0.33588019058041785,
'n_estimators': 242, 'min_child_weight': 4, 'subsample': 0.8762844269722132,
'colsample_bytree': 0.549676384981236, 'gamma': 7.563888272016332e-06,
'reg_alpha': 0.012499114906749 43, 'reg_lambda': 1.5195258576188784e-06}. Best is
trial 19 with value: 0.0012202559034307063.
[I 2024-07-14 16:20:46,919] Trial 29 finished with value: 0.010235079833469 06
and parameters: {'max_depth': 3, 'learning_rate': 0.045788446958940465,
'n_estimators': 547, 'min_child_weight': 2, 'subsample': 0.7764048708096791,

'colsample_bytree': 0.7544838593841495, 'gamma': 1.7982726202518847e-07,
'reg_alpha': 0.00011240013673597403, 'reg_lambda': 0.00019791808680946632}. Best
is trial 19 with value: 0.0012202559034307063.
[I 2024-07-14 16:20:49,785] Trial 30 finished with value: 0.010791710257084785
and parameters: {'max_depth': 5, 'learning_rate': 0.6054683181491443,
'n_estimators': 841, 'min_child_weight': 2, 'subsample': 0.9680435839407016,
'colsample_bytree': 0.6265185122256889, 'gamma': 1.9517909027504355e-06,
'reg_alpha': 3.887043140294087e-06, 'reg_lambda': 0.0018240221885705045}. Best
is trial 19 with value: 0.0012202559034307063.
[I 2024-07-14 16:20:52,006] Trial 31 finished with value: 0.004638899926257148
and parameters: {'max_depth': 3, 'learning_rate': 0.6318475060872858,
'n_estimators': 957, 'min_child_weight': 1, 'subsample': 0.9111873683070938,
'colsample_bytree': 0.6712841538847817, 'gamma': 1.0275193657627934e-05,
'reg_alpha': 5.0775481185341695e-05, 'reg_lambda': 1.8244776722888047e-05}. Best
is trial 19 with value: 0.0012202559034307063.
[I 2024-07-14 16:20:53,756] Trial 32 finished with value: 0.0039977347594007725
and parameters: {'max_depth': 2, 'learning_rate': 0.16050632370695275,
'n_estimators': 995, 'min_child_weight': 1, 'subsample': 0.9230700259109924,
'colsample_bytree': 0.7729791392986893, 'gamma': 5.802043475250585e-06,
'reg_alpha': 5.382289454219401e-05, 'reg_lambda': 3.9508597703550035e-06}. Best
is trial 19 with value: 0.0012202559034307063.
[I 2024-07-14 16:20:56,155] Trial 33 finished with value: 0.0029980551576832275
and parameters: {'max_depth': 2, 'learning_rate': 0.5948287083191321,
'n_estimators': 1200, 'min_child_weight': 3, 'subsample': 0.8495904038181203,
'colsample_bytree': 0.7751387923499603, 'gamma': 1.0361631905866391e-05,
'reg_alpha': 5.2348920252052974e-05, 'reg_lambda': 5.667963055703207e-06}. Best
is trial 19 with value: 0.0012202559034307063.
[I 2024-07-14 16:20:57,985] Trial 34 finished with value: 0.0024737328888596987
and parameters: {'max_depth': 1, 'learning_rate': 0.603327836280234,
'n_estimators': 1192, 'min_child_weight': 4, 'subsample': 0.844172488239402,
'colsample_bytree': 0.769480104123524, 'gamma': 9.519083331790536e-06,
'reg_alpha': 0.00031013107916128007, 'reg_lambda': 6.254001140612522e-06}. Best
is trial 19 with value: 0.0012202559034307063.
[I 2024-07-14 16:20:59,847] Trial 35 finished with value: 0.0012989134828420108
and parameters: {'max_depth': 1, 'learning_rate': 0.39315311132114283,
'n_estimators': 1222, 'min_child_weight': 4, 'subsample': 0.8293078354581684,
'colsample_bytree': 0.7842311633695939, 'gamma': 0.0001571577861828617,
'reg_alpha': 0.0003543784644747976, 'reg_lambda': 3.5094434653369267e-06}. Best
is trial 19 with value: 0.0012202559034307063.
[I 2024-07-14 16:21:01,589] Trial 36 finished with value: 0.0011901107640401002
and parameters: {'max_depth': 1, 'learning_rate': 0.3467969192122621,
'n_estimators': 1215, 'min_child_weight': 5, 'subsample': 0.841077268420862,
'colsample_bytree': 0.8983709637976001, 'gamma': 0.00015961795650361657,
'reg_alpha': 0.000263557241115935, 'reg_lambda': 1.493010779391425e-07}. Best is
trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:03,804] Trial 37 finished with value: 0.0045841525361867514
and parameters: {'max_depth': 1, 'learning_rate': 0.38623302550041994,
'n_estimators': 1640, 'min_child_weight': 5, 'subsample': 0.7134199949062017,

'colsample_bytree': 0.9948385281020939, 'gamma': 0.00655345179122262,
'reg_alpha': 0.0004256533237882259, 'reg_lambda': 1.9577918737670465e-07}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:06,160] Trial 38 finished with value: 0.022198009431117684
and parameters: {'max_depth': 12, 'learning_rate': 0.33317267397745787,
'n_estimators': 1535, 'min_child_weight': 6, 'subsample': 0.8277374102038245,
'colsample_bytree': 0.9059540480686992, 'gamma': 0.000211878592929239,
'reg_alpha': 0.0002516077039563867, 'reg_lambda': 7.648276253079233e-08}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:09,252] Trial 39 finished with value: 0.01687058934604575
and parameters: {'max_depth': 17, 'learning_rate': 0.09013420204093052,
'n_estimators': 1270, 'min_child_weight': 6, 'subsample': 0.7722022876631909,
'colsample_bytree': 0.9446398054622651, 'gamma': 0.000811974933696692,
'reg_alpha': 0.01758847281971534, 'reg_lambda': 3.103985732660924e-07}. Best is
trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:16,423] Trial 40 finished with value: 0.006672466814517014
and parameters: {'max_depth': 6, 'learning_rate': 0.019173959468839333,
'n_estimators': 1789, 'min_child_weight': 4, 'subsample': 0.8489981290151112,
'colsample_bytree': 0.9096169504813136, 'gamma': 3.6698638740750816e-05,
'reg_alpha': 4.160742964322052e-06, 'reg_lambda': 9.051669407595633e-05}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:18,080] Trial 41 finished with value: 0.0025147312186199486
and parameters: {'max_depth': 1, 'learning_rate': 0.5711448581667254,
'n_estimators': 1174, 'min_child_weight': 4, 'subsample': 0.8624393626556283,
'colsample_bytree': 0.7859800414271598, 'gamma': 9.324498923837512e-05,
'reg_alpha': 7.417701631886109e-05, 'reg_lambda': 4.215183092421243e-06}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:19,786] Trial 42 finished with value: 0.002437536971097205
and parameters: {'max_depth': 1, 'learning_rate': 0.7048718072385178,
'n_estimators': 1110, 'min_child_weight': 4, 'subsample': 0.874359916067545,
'colsample_bytree': 0.8130793156640406, 'gamma': 0.00012983217465028488,
'reg_alpha': 0.00019001501008819078, 'reg_lambda': 5.768697024305781e-07}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:21,170] Trial 43 finished with value: 0.009847593571958604
and parameters: {'max_depth': 2, 'learning_rate': 0.8314536326084561,
'n_estimators': 1055, 'min_child_weight': 5, 'subsample': 0.8811981774450696,
'colsample_bytree': 0.8460362968580881, 'gamma': 0.0034888782577876883,
'reg_alpha': 0.00027036581162739247, 'reg_lambda': 4.610588821665775e-08}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:26,174] Trial 44 finished with value: 0.03972191210867478
and parameters: {'max_depth': 4, 'learning_rate': 0.0014245082651810324,
'n_estimators': 1325, 'min_child_weight': 5, 'subsample': 0.8013052275920441,
'colsample_bytree': 0.8198388104818279, 'gamma': 3.139643531230183e-05,
'reg_alpha': 0.0005544854007609853, 'reg_lambda': 4.5215568959767357e-07}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:27,894] Trial 45 finished with value: 0.002266848311382039
and parameters: {'max_depth': 1, 'learning_rate': 0.48569789425980575,
'n_estimators': 1107, 'min_child_weight': 4, 'subsample': 0.8280580135328051,

'colsample_bytree': 0.7136225891949566, 'gamma': 0.0003150404907236029,
'reg_alpha': 7.670458816667178e-06, 'reg_lambda': 7.028522334776742e-07}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:30,147] Trial 46 finished with value: 0.01877305517808857
and parameters: {'max_depth': 20, 'learning_rate': 0.4390435954607433,
'n_estimators': 1431, 'min_child_weight': 6, 'subsample': 0.7582253819096745,
'colsample_bytree': 0.7321660172623896, 'gamma': 0.027600793976666513,
'reg_alpha': 1.1217529152777647e-06, 'reg_lambda': 3.9730200176099255e-07}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:33,897] Trial 47 finished with value: 0.02646123992684531
and parameters: {'max_depth': 4, 'learning_rate': 0.0054121783545049455,
'n_estimators': 1097, 'min_child_weight': 4, 'subsample': 0.7263332148569077,
'colsample_bytree': 0.7071613751998567, 'gamma': 0.00033231958884567315,
'reg_alpha': 9.010516308572383e-06, 'reg_lambda': 4.934470467501106e-08}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:35,191] Trial 48 finished with value: 0.005696929098868925
and parameters: {'max_depth': 2, 'learning_rate': 0.2881318181823724,
'n_estimators': 752, 'min_child_weight': 5, 'subsample': 0.6847103725364925,
'colsample_bytree': 0.8159611615118209, 'gamma': 0.0006048874434217591,
'reg_alpha': 1.4586579937243174e-06, 'reg_lambda': 6.054247915242309e-07}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:36,753] Trial 49 finished with value: 0.0024108771395600277
and parameters: {'max_depth': 1, 'learning_rate': 0.7349690174300101,
'n_estimators': 1050, 'min_child_weight': 3, 'subsample': 0.8220191091934597,
'colsample_bytree': 0.8824384489501733, 'gamma': 0.0001613565646796602,
'reg_alpha': 5.228853567080754e-07, 'reg_lambda': 2.665582122678115e-06}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:39,104] Trial 50 finished with value: 0.014979244694982932
and parameters: {'max_depth': 15, 'learning_rate': 0.12193072881338855,
'n_estimators': 1020, 'min_child_weight': 8, 'subsample': 0.5581156243093073,
'colsample_bytree': 0.8923425775749741, 'gamma': 0.0009001243759486723,
'reg_alpha': 3.9066873498806807e-07, 'reg_lambda': 2.120348898468363e-06}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:41,029] Trial 51 finished with value: 0.0019255573425601552
and parameters: {'max_depth': 1, 'learning_rate': 0.4724220590974576,
'n_estimators': 1254, 'min_child_weight': 3, 'subsample': 0.8236087981338818,
'colsample_bytree': 0.8656070769135761, 'gamma': 0.00014206908700707357,
'reg_alpha': 5.834302586177408e-06, 'reg_lambda': 1.2602438913846857e-05}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:42,947] Trial 52 finished with value: 0.004830021163517983
and parameters: {'max_depth': 2, 'learning_rate': 0.4475489605405081,
'n_estimators': 1246, 'min_child_weight': 3, 'subsample': 0.8190887567610701,
'colsample_bytree': 0.8534989267730576, 'gamma': 0.00022053735688833882,
'reg_alpha': 5.163899322323676e-06, 'reg_lambda': 1.1541043230245955e-05}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:46,051] Trial 53 finished with value: 0.005563109936702931
and parameters: {'max_depth': 1, 'learning_rate': 0.8037701271784371,
'n_estimators': 1464, 'min_child_weight': 4, 'subsample': 0.827062779111593,

'colsample_bytree': 0.8802701829979992, 'gamma': 0.0038657233957332484,
'reg_alpha': 4.867646440124573e-07, 'reg_lambda': 4.7776611940129515e-05}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:50,936] Trial 54 finished with value: 0.008270179296709923
and parameters: {'max_depth': 5, 'learning_rate': 0.2152707262991497,
'n_estimators': 1310, 'min_child_weight': 3, 'subsample': 0.7897999962504614,
'colsample_bytree': 0.9330831995690371, 'gamma': 6.508917744591784e-05,
'reg_alpha': 7.131359441034996e-08, 'reg_lambda': 2.5721934137921933e-06}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:55,439] Trial 55 finished with value: 0.005549062774882251
and parameters: {'max_depth': 4, 'learning_rate': 0.3048490216379295,
'n_estimators': 1557, 'min_child_weight': 2, 'subsample': 0.7544797047161763,
'colsample_bytree': 0.9716665987670513, 'gamma': 2.5455503756347152e-05,
'reg_alpha': 7.380259016330037e-06, 'reg_lambda': 1.29140226893074e-05}. Best is
trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:57,824] Trial 56 finished with value: 0.03840986427009843
and parameters: {'max_depth': 3, 'learning_rate': 0.003572559758836438,
'n_estimators': 831, 'min_child_weight': 3, 'subsample': 0.8932720934810893,
'colsample_bytree': 0.9185662039943403, 'gamma': 0.00044231176892039424,
'reg_alpha': 5.155912128040879e-08, 'reg_lambda': 2.2931951964568467e-08}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:21:59,899] Trial 57 finished with value: 0.004921674802448375
and parameters: {'max_depth': 2, 'learning_rate': 0.44931575348465624,
'n_estimators': 1375, 'min_child_weight': 4, 'subsample': 0.8097145723085752,
'colsample_bytree': 0.73872648755336, 'gamma': 0.00018392197823878309,
'reg_alpha': 2.1629920134166737e-06, 'reg_lambda': 1.307025490346696e-07}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:01,628] Trial 58 finished with value: 0.03853486040819823
and parameters: {'max_depth': 1, 'learning_rate': 0.007826085424791816,
'n_estimators': 1131, 'min_child_weight': 2, 'subsample': 0.8287628022928661,
'colsample_bytree': 0.6487341345692291, 'gamma': 0.001280360165490179,
'reg_alpha': 5.03286567760455e-07, 'reg_lambda': 9.652863656222883e-05}. Best is
trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:03,985] Trial 59 finished with value: 0.0070773121281768655
and parameters: {'max_depth': 3, 'learning_rate': 0.9353350469810336,
'n_estimators': 1008, 'min_child_weight': 3, 'subsample': 0.7744831125138597,
'colsample_bytree': 0.8573611365479292, 'gamma': 3.0552727934015674e-06,
'reg_alpha': 1.05502286659453e-05, 'reg_lambda': 2.98469059906973e-06}. Best is
trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:05,662] Trial 60 finished with value: 0.0024111086934797115
and parameters: {'max_depth': 1, 'learning_rate': 0.2541300983272984,
'n_estimators': 1238, 'min_child_weight': 5, 'subsample': 0.7974594030268742,
'colsample_bytree': 0.8342631562747791, 'gamma': 0.002948099262213289,
'reg_alpha': 3.345556051433405e-05, 'reg_lambda': 8.524756753966865e-06}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:07,703] Trial 61 finished with value: 0.0051995846237762305
and parameters: {'max_depth': 1, 'learning_rate': 0.2832890795404859,
'n_estimators': 1253, 'min_child_weight': 5, 'subsample': 0.7982624071767797,

'colsample_bytree': 0.8828703675808314, 'gamma': 0.01505983636887358,
'reg_alpha': 3.513816115869352e-05, 'reg_lambda': 1.1058084719604184e-05}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:09,226] Trial 62 finished with value: 0.00583265889871925
and parameters: {'max_depth': 2, 'learning_rate': 0.2245426195233497,
'n_estimators': 1059, 'min_child_weight': 6, 'subsample': 0.8502334511742563,
'colsample_bytree': 0.8332212544447702, 'gamma': 0.0022798147855370944,
'reg_alpha': 0.00010841368614503232, 'reg_lambda': 8.629906369372055e-07}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:11,172] Trial 63 finished with value: 0.0020782608717838658
and parameters: {'max_depth': 1, 'learning_rate': 0.48769968279705733,
'n_estimators': 1220, 'min_child_weight': 5, 'subsample': 0.7390982495914632,
'colsample_bytree': 0.8924692507606724, 'gamma': 0.00010020443027103781,
'reg_alpha': 2.190074021037009e-06, 'reg_lambda': 7.437749560154959e-06}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:13,622] Trial 64 finished with value: 0.007933803076241114
and parameters: {'max_depth': 4, 'learning_rate': 0.4948863441909777,
'n_estimators': 904, 'min_child_weight': 4, 'subsample': 0.6560311242663746,
'colsample_bytree': 0.8915090441227286, 'gamma': 0.00010096946238085114,
'reg_alpha': 9.841651231081693e-07, 'reg_lambda': 2.5012421634154694e-05}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:15,367] Trial 65 finished with value: 0.0038845033746300744
and parameters: {'max_depth': 2, 'learning_rate': 0.7276073391946518,
'n_estimators': 969, 'min_child_weight': 3, 'subsample': 0.7423238000296106,
'colsample_bytree': 0.8641967770160512, 'gamma': 5.016476235513998e-05,
'reg_alpha': 2.433895754119226e-06, 'reg_lambda': 4.34448926747581e-05}. Best is
trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:18,586] Trial 66 finished with value: 0.003867501216877539
and parameters: {'max_depth': 3, 'learning_rate': 0.3620653365015708,
'n_estimators': 1313, 'min_child_weight': 5, 'subsample': 0.6985821095037003,
'colsample_bytree': 0.7972130602696252, 'gamma': 1.659408291354708e-05,
'reg_alpha': 1.2741987919433107e-05, 'reg_lambda': 2.0243928983643804e-06}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:20,308] Trial 67 finished with value: 0.0022692948303684895
and parameters: {'max_depth': 1, 'learning_rate': 0.4893291534647043,
'n_estimators': 1153, 'min_child_weight': 3, 'subsample': 0.7390583211333891,
'colsample_bytree': 0.965405386186531, 'gamma': 0.00031539186064819067,
'reg_alpha': 1.5198253027538315e-07, 'reg_lambda': 0.00014906991469394121}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:22,302] Trial 68 finished with value: 0.0047178583837596145
and parameters: {'max_depth': 2, 'learning_rate': 0.18443622566713502,
'n_estimators': 1391, 'min_child_weight': 4, 'subsample': 0.9693946717414226,
'colsample_bytree': 0.976620244767594, 'gamma': 0.0003433909811159869,
'reg_alpha': 1.0965538304337345e-08, 'reg_lambda': 0.00048785126213516533}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:25,396] Trial 69 finished with value: 0.01443233337012483
and parameters: {'max_depth': 11, 'learning_rate': 0.067461861590224,
'n_estimators': 1119, 'min_child_weight': 2, 'subsample': 0.6381384981892587,

'colsample_bytree': 0.9616168550634046, 'gamma': 0.0007310261681741371,
'reg_alpha': 1.9854703937593585e-07, 'reg_lambda': 0.00019863191093079139}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:28,078] Trial 70 finished with value: 0.00747544609390626
and parameters: {'max_depth': 5, 'learning_rate': 0.1338325529912911,
'n_estimators': 1211, 'min_child_weight': 4, 'subsample': 0.7327088734287608,
'colsample_bytree': 0.6895736739474322, 'gamma': 9.891734321739468e-05,
'reg_alpha': 0.00095156407589836, 'reg_lambda': 8.21228634065309e-05}. Best is
trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:29,901] Trial 71 finished with value: 0.0013338107800616049
and parameters: {'max_depth': 1, 'learning_rate': 0.38866711231783235,
'n_estimators': 1147, 'min_child_weight': 3, 'subsample': 0.7729771543395406,
'colsample_bytree': 0.9237578399568787, 'gamma': 0.00013315118773123084,
'reg_alpha': 1.0117715724808372e-07, 'reg_lambda': 1.0141917566622787e-06}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:31,687] Trial 72 finished with value: 0.0016401544975582316
and parameters: {'max_depth': 1, 'learning_rate': 0.3986722706466829,
'n_estimators': 1157, 'min_child_weight': 3, 'subsample': 0.7830132019098375,
'colsample_bytree': 0.9322849422867865, 'gamma': 0.00028047029776879996,
'reg_alpha': 8.441697861060791e-08, 'reg_lambda': 1.1027703174654762e-06}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:34,190] Trial 73 finished with value: 0.003427802252488179
and parameters: {'max_depth': 2, 'learning_rate': 0.3799562179163471,
'n_estimators': 1310, 'min_child_weight': 2, 'subsample': 0.7719605509053964,
'colsample_bytree': 0.9254343401400743, 'gamma': 5.0396884896911615e-05,
'reg_alpha': 3.266767105175015e-08, 'reg_lambda': 9.747236804848032e-07}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:36,752] Trial 74 finished with value: 0.0016126944979776626
and parameters: {'max_depth': 1, 'learning_rate': 0.31599453810265943,
'n_estimators': 1979, 'min_child_weight': 3, 'subsample': 0.7595775822245425,
'colsample_bytree': 0.9424471971142012, 'gamma': 0.0004853889143204719,
'reg_alpha': 7.789902716337268e-08, 'reg_lambda': 1.2857523397996807e-07}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:41,341] Trial 75 finished with value: 0.0035256154567230333
and parameters: {'max_depth': 3, 'learning_rate': 0.19055247311616322,
'n_estimators': 1982, 'min_child_weight': 3, 'subsample': 0.7585068633144854,
'colsample_bytree': 0.9399290819659116, 'gamma': 6.384004471951746e-07,
'reg_alpha': 9.960563417752475e-08, 'reg_lambda': 2.146381485150578e-07}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:44,597] Trial 76 finished with value: 0.0027362885407797087
and parameters: {'max_depth': 2, 'learning_rate': 0.31625093992582387,
'n_estimators': 1761, 'min_child_weight': 2, 'subsample': 0.7155926165729734,
'colsample_bytree': 0.9088774922507689, 'gamma': 2.3300835341901086e-05,
'reg_alpha': 2.296847371201515e-08, 'reg_lambda': 1.0961445332471431e-07}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:46,770] Trial 77 finished with value: 0.0028240907174934466
and parameters: {'max_depth': 1, 'learning_rate': 0.10197308878910082,
'n_estimators': 1552, 'min_child_weight': 3, 'subsample': 0.786523477609062,

'colsample_bytree': 0.9514599731713888, 'gamma': 0.000558182394632845,
'reg_alpha': 2.615595339031453e-07, 'reg_lambda': 6.120440656345671e-06}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:48,223] Trial 78 finished with value: 0.006523897027397389
and parameters: {'max_depth': 3, 'learning_rate': 0.15653578986684105,
'n_estimators': 755, 'min_child_weight': 10, 'subsample': 0.7036815639052468,
'colsample_bytree': 0.9959467979615902, 'gamma': 0.0012421677014272593,
'reg_alpha': 5.141658699696563e-08, 'reg_lambda': 2.3997486407867704e-07}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:51,649] Trial 79 finished with value: 0.01798353963406139
and parameters: {'max_depth': 8, 'learning_rate': 0.386526476666489,
'n_estimators': 1855, 'min_child_weight': 4, 'subsample': 0.7639164237299152,
'colsample_bytree': 0.9218605022347864, 'gamma': 9.650685537140057e-05,
'reg_alpha': 1.000730157090367e-07, 'reg_lambda': 3.057830563492324e-08}. Best
is trial 36 with value: 0.0011901107640401002.
[I 2024-07-14 16:22:54,006] Trial 80 finished with value: 0.0005090415916422417
and parameters: {'max_depth': 1, 'learning_rate': 0.2464355735520665,
'n_estimators': 1691, 'min_child_weight': 3, 'subsample': 0.8612364164070719,
'colsample_bytree': 0.8980411988917243, 'gamma': 4.0152994870799015e-05,
'reg_alpha': 2.4155799800674818e-08, 'reg_lambda': 1.3182273975909007e-06}. Best
is trial 80 with value: 0.0005090415916422417.
[I 2024-07-14 16:22:56,690] Trial 81 finished with value: 0.0005224549383774007
and parameters: {'max_depth': 1, 'learning_rate': 0.24516918402898533,
'n_estimators': 1959, 'min_child_weight': 3, 'subsample': 0.8645562176343059,
'colsample_bytree': 0.8992468143249244, 'gamma': 3.65023704536492e-05,
'reg_alpha': 2.6436281972715583e-08, 'reg_lambda': 1.4430913201143605e-06}. Best
is trial 80 with value: 0.0005090415916422417.
[I 2024-07-14 16:23:00,033] Trial 82 finished with value: 0.0028261843548368876
and parameters: {'max_depth': 2, 'learning_rate': 0.20107645111831254,
'n_estimators': 1970, 'min_child_weight': 3, 'subsample': 0.8853001785509027,
'colsample_bytree': 0.899908523096414, 'gamma': 1.2720068209340486e-05,
'reg_alpha': 3.165845347974043e-08, 'reg_lambda': 1.0753031384774945e-06}. Best
is trial 80 with value: 0.0005090415916422417.
[I 2024-07-14 16:23:02,465] Trial 83 finished with value: 0.0004691860664770545
and parameters: {'max_depth': 1, 'learning_rate': 0.2411982169515882,
'n_estimators': 1662, 'min_child_weight': 2, 'subsample': 0.8608017749518633,
'colsample_bytree': 0.9546650086425718, 'gamma': 2.74680987895295e-05,
'reg_alpha': 1.4255563568433152e-08, 'reg_lambda': 1.4071932469410387e-06}. Best
is trial 83 with value: 0.0004691860664770545.
[I 2024-07-14 16:23:06,186] Trial 84 finished with value: 0.003933964357896243
and parameters: {'max_depth': 3, 'learning_rate': 0.2653454983829556,
'n_estimators': 1911, 'min_child_weight': 2, 'subsample': 0.8608847324973327,
'colsample_bytree': 0.9841829778606073, 'gamma': 3.113543141292643e-05,
'reg_alpha': 1.430281800584647e-08, 'reg_lambda': 1.5275187436594104e-06}. Best
is trial 83 with value: 0.0004691860664770545.
[I 2024-07-14 16:23:09,140] Trial 85 finished with value: 0.003236421411564326
and parameters: {'max_depth': 2, 'learning_rate': 0.15539797165006827,
'n_estimators': 1697, 'min_child_weight': 2, 'subsample': 0.9187750299055781,

'colsample_bytree': 0.9480095143621896, 'gamma': 6.599280117231726e-06,
'reg_alpha': 2.1971426965173032e-08, 'reg_lambda': 5.054640620413229e-07}. Best
is trial 83 with value: 0.0004691860664770545.
[I 2024-07-14 16:23:19,867] Trial 86 finished with value: 0.0223151423789554 and
parameters: {'max_depth': 18, 'learning_rate': 0.05011663093599488,
'n_estimators': 1828, 'min_child_weight': 2, 'subsample': 0.8403168435365238,
'colsample_bytree': 0.956802336985689, 'gamma': 4.407407530881073e-05,
'reg_alpha': 1.1313889456640058e-08, 'reg_lambda': 7.63827620013783e-08}. Best
is trial 83 with value: 0.0004691860664770545.
[I 2024-07-14 16:23:22,199] Trial 87 finished with value: 0.0004495217920432591
and parameters: {'max_depth': 1, 'learning_rate': 0.22990165322641637,
'n_estimators': 1656, 'min_child_weight': 3, 'subsample': 0.8607280681867803,
'colsample_bytree': 0.9333231657847741, 'gamma': 3.0788302510592997e-06,
'reg_alpha': 3.2733517312924784e-08, 'reg_lambda': 3.6781446658299294e-07}. Best
is trial 87 with value: 0.0004495217920432591.
[I 2024-07-14 16:23:28,199] Trial 88 finished with value: 0.020723383610207466
and parameters: {'max_depth': 14, 'learning_rate': 0.08757006254772103,
'n_estimators': 1606, 'min_child_weight': 3, 'subsample': 0.8649077666459111,
'colsample_bytree': 0.9155013048913974, 'gamma': 4.12047390908931e-06,
'reg_alpha': 3.497850189595827e-08, 'reg_lambda': 3.191514964336707e-07}. Best
is trial 87 with value: 0.0004495217920432591.
[I 2024-07-14 16:23:31,275] Trial 89 finished with value: 0.002932163727479713
and parameters: {'max_depth': 2, 'learning_rate': 0.2450697468826984,
'n_estimators': 1721, 'min_child_weight': 3, 'subsample': 0.889143245528718,
'colsample_bytree': 0.6560229191090665, 'gamma': 1.10158145073574e-06,
'reg_alpha': 1.8914488806236925e-08, 'reg_lambda': 0.24110119616109957}. Best is
trial 87 with value: 0.0004495217920432591.
[I 2024-07-14 16:23:36,277] Trial 90 finished with value: 0.00511396170682232
and parameters: {'max_depth': 4, 'learning_rate': 0.025097084690598917,
'n_estimators': 1899, 'min_child_weight': 4, 'subsample': 0.9297178256729757,
'colsample_bytree': 0.9377659456259974, 'gamma': 1.5320563151920318e-06,
'reg_alpha': 0.002586629149914586, 'reg_lambda': 1.3821574243959887e-07}. Best
is trial 87 with value: 0.0004495217920432591.
[I 2024-07-14 16:23:38,730] Trial 91 finished with value: 0.0006699100789707718
and parameters: {'max_depth': 1, 'learning_rate': 0.3269575420532919,
'n_estimators': 1652, 'min_child_weight': 3, 'subsample': 0.9073155857404581,
'colsample_bytree': 0.9329432583462608, 'gamma': 1.379898376689046e-05,
'reg_alpha': 7.784837322262208e-08, 'reg_lambda': 1.6817500337428371e-06}. Best
is trial 87 with value: 0.0004495217920432591.
[I 2024-07-14 16:23:41,849] Trial 92 finished with value: 0.0005019289954238832
and parameters: {'max_depth': 1, 'learning_rate': 0.3007257694138011,
'n_estimators': 1942, 'min_child_weight': 3, 'subsample': 0.838620338962275,
'colsample_bytree': 0.9790179068260536, 'gamma': 2.630671387248035e-06,
'reg_alpha': 5.691023037382993e-08, 'reg_lambda': 1.6555069263189744e-06}. Best
is trial 87 with value: 0.0004495217920432591.
[I 2024-07-14 16:23:44,020] Trial 93 finished with value: 0.0014842610400818254
and parameters: {'max_depth': 1, 'learning_rate': 0.13099576578257446,
'n_estimators': 1609, 'min_child_weight': 2, 'subsample': 0.9005629444530807,

```
'colsample_bytree': 0.98168358553467, 'gamma': 1.2963671479747432e-05,
'reg_alpha': 4.241565695538955e-08, 'reg_lambda': 3.57914614062562e-06}. Best is
trial 87 with value: 0.0004495217920432591.
[I 2024-07-14 16:23:46,377] Trial 94 finished with value: 0.0004922484115284476
and parameters: {'max_depth': 1, 'learning_rate': 0.22519519241472966,
'n_estimators': 1646, 'min_child_weight': 4, 'subsample': 0.8582997382628113,
'colsample_bytree': 0.9646965680506456, 'gamma': 2.8221298590857956e-06,
'reg_alpha': 2.349504423926681e-08, 'reg_lambda': 1.8330286636334817e-06}. Best
is trial 87 with value: 0.0004495217920432591.
[I 2024-07-14 16:23:49,746] Trial 95 finished with value: 0.002937721295775006
and parameters: {'max_depth': 2, 'learning_rate': 0.21502387242950055,
'n_estimators': 1747, 'min_child_weight': 4, 'subsample': 0.9095753008515335,
'colsample_bytree': 0.9998841781664333, 'gamma': 3.1688396120193557e-06,
'reg_alpha': 2.4181523489574596e-08, 'reg_lambda': 1.7524494897104383e-06}. Best
is trial 87 with value: 0.0004495217920432591.
[I 2024-07-14 16:23:52,007] Trial 96 finished with value: 0.000356842852669995
and parameters: {'max_depth': 1, 'learning_rate': 0.269815464158218,
'n_estimators': 1658, 'min_child_weight': 1, 'subsample': 0.8601935079349704,
'colsample_bytree': 0.5048453403830055, 'gamma': 2.1943345278265024e-06,
'reg_alpha': 1.4724579285012502e-08, 'reg_lambda': 3.9200215686546185e-06}. Best
is trial 96 with value: 0.000356842852669995.
[I 2024-07-14 16:23:55,742] Trial 97 finished with value: 0.0042030916650765 and
parameters: {'max_depth': 3, 'learning_rate': 0.18202654730693293,
'n_estimators': 1669, 'min_child_weight': 1, 'subsample': 0.9595244631986798,
'colsample_bytree': 0.5008105028754477, 'gamma': 4.507164716709046e-07,
'reg_alpha': 1.539470595709278e-08, 'reg_lambda': 6.271929109956166e-07}. Best
is trial 96 with value: 0.000356842852669995.
[I 2024-07-14 16:23:58,447] Trial 98 finished with value: 0.0031100352787020932
and parameters: {'max_depth': 2, 'learning_rate': 0.26654674105764403,
'n_estimators': 1485, 'min_child_weight': 1, 'subsample': 0.8693035080722711,
'colsample_bytree': 0.5471374371103065, 'gamma': 2.4914434896672427e-06,
'reg_alpha': 5.682807933567933e-08, 'reg_lambda': 4.2974552465184e-06}. Best is
trial 96 with value: 0.000356842852669995.
[I 2024-07-14 16:24:01,030] Trial 99 finished with value: 0.0006363583382324991
and parameters: {'max_depth': 1, 'learning_rate': 0.16658028527204613,
'n_estimators': 1813, 'min_child_weight': 2, 'subsample': 0.8527566230022827,
'colsample_bytree': 0.9877277687158535, 'gamma': 2.490583055628943e-07,
'reg_alpha': 2.6725926199650874e-08, 'reg_lambda': 3.4552635306048054e-07}. Best
is trial 96 with value: 0.000356842852669995.
```

[19]:
```
best_params = study.best_params
```

[20]:
```
model = build_model(best_params)
model.fit(X_train, y_train)
predictions = model.predict(X_test)
print_metrics(y_test, predictions)
```

```
Root Mean Squarred Error: 0.00026314685904943897
```

```
R-squared (r^2): 0.9999720505723292
Mean Absolute Error: 0.0001825627629756932
```

[ ]: