

# model1

November 27, 2024

## 1 Model 1

### 1.1 KNN -> K-means -> GOSDT on Individual Clusters

```
[ ]: import pandas as pd
import numpy as np
import sklearn as sk
import math
from sklearn.neighbors import NearestNeighbors
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans
from sklearn.decomposition import PCA
from sklearn.model_selection import GridSearchCV
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import classification_report
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import MinMaxScaler
from sklearn.preprocessing import StandardScaler
from sklearn.metrics.pairwise import euclidean_distances
import matplotlib.pyplot as plt
import numpy as np
from sklearn.model_selection import KFold, ParameterGrid
from sklearn.metrics import mean_squared_error
import gosdt
```

### 1.2 Merge the train and test datasets

```
[ ]: pd.set_option('display.max_columns', None)

# Load the dfs
train_df = pd.read_csv('train_processed.csv')
test_df = pd.read_csv('test_processed.csv')

# Find the lengths of the train and test dataframes
train_size = len(train_df)
test_size = len(test_df)

# combine them for unsupervised learning
```

```
combined_df = pd.concat([train_df, test_df], axis=0)

# save the price and id columns for later
price_column = combined_df['price']
id_column = combined_df['id']

combined_df = combined_df.drop(columns=['price', 'id'])

combined_df = combined_df.reset_index(drop=True)
price_column = price_column.reset_index(drop=True)
id_column = id_column.reset_index(drop=True)

combined_df
```

```
[ ]:      latitude  longitude  host_since  host_response_time  \
0      40.684560 -73.939870      16578           2.0
1      40.638991 -73.965739      19614           1.0
2      40.618810 -74.032380      19204           1.0
3      40.673970 -73.953990      15563           1.0
4      40.747180 -73.985390      16427           1.0
...      ...      ...      ...      ...
22418  40.637960 -73.951360      16690           1.0
22419  40.823720 -73.945460      16877           4.0
22420  40.755094 -73.937260      15563           1.0
22421  40.781580 -73.984780      15351           1.0
22422  40.695890 -73.760830      17286           1.0

      host_response_rate  host_acceptance_rate  host_is_superhost  \
0              100.0              100.0              1
1              100.0              98.0              1
2              100.0              100.0              0
3              99.0              23.0              0
4              93.0              95.0              0
...      ...      ...      ...
22418          100.0              100.0              0
22419           0.0              33.0              0
22420          99.0              23.0              0
22421          100.0              100.0              1
22422          100.0              100.0              0

      host_listings_count  host_total_listings_count  host_has_profile_pic  \
0              2.0              2.0              1
1              1.0              1.0              1
2              52.0              55.0              1
3             727.0             1336.0              1
4             707.0             2453.0              1
...      ...      ...      ...
```

22418	2.0	2.0	1
22419	7.0	8.0	1
22420	727.0	1336.0	1
22421	1.0	3.0	1
22422	3.0	3.0	1

	host_identity_verified	calculated_host_listings_count	accommodates	\
0	1	1	4	
1	1	1	2	
2	1	52	2	
3	1	719	1	
4	1	73	2	
...	...	...	...	
22418	1	2	2	
22419	0	7	1	
22420	1	719	1	
22421	1	1	2	
22422	1	3	2	

	bathrooms	bedrooms	beds	has_availability	availability_30	\
0	2.0	2.0	2.0	1	12	
1	1.0	1.0	2.0	1	10	
2	1.0	0.0	1.0	1	17	
3	1.5	4.0	1.0	1	0	
4	1.0	1.0	1.0	1	4	
...	...	...	...	...	...	
22418	1.0	2.0	6.0	1	30	
22419	1.0	1.0	1.0	1	0	
22420	3.0	5.0	1.0	1	3	
22421	1.0	1.0	1.0	1	18	
22422	1.0	1.0	1.0	1	30	

	availability_60	availability_90	availability_365	instant_bookable	\
0	42	70	70	False	
1	20	49	324	False	
2	44	70	146	True	
3	0	0	111	False	
4	13	22	241	True	
...	...	...	...	...	
22418	60	90	365	False	
22419	1	1	97	False	
22420	33	63	159	False	
22421	35	35	35	False	
22422	60	90	180	False	

	minimum_nights	maximum_nights	number_of_reviews	\
0	30	1125	34	

1	1	29	30
2	1	29	5
3	30	365	0
4	1	1125	0
...	...	...	...
22418	30	1125	325
22419	30	365	11
22420	30	365	0
22421	30	30	33
22422	30	90	19

	number_of_reviews_ltm	number_of_reviews_l30d	first_review \
0	5	1	18014
1	30	6	19735
2	5	2	19901
3	0	0	20049
4	0	0	20049
...	...	...	...
22418	9	0	16729
22419	2	0	17774
22420	0	0	20049
22421	6	0	18962
22422	2	0	19121

	last_review	review_scores_rating	review_scores_accuracy \
0	19945	5.000000	5.000000
1	19968	4.830000	4.870000
2	19952	4.600000	4.800000
3	20049	4.719393	4.742812
4	20049	4.719393	4.742812
...	...	...	...
22418	19679	4.710000	4.750000
22419	19875	4.360000	4.550000
22420	20049	4.711449	4.748402
22421	19906	4.880000	4.940000
22422	19813	4.740000	4.680000

	review_scores_cleanliness	review_scores_checkin \
0	4.970000	5.000000
1	4.930000	4.800000
2	4.200000	4.800000
3	4.679642	4.826310
4	4.679642	4.826310
...	...	...
22418	4.640000	4.880000
22419	4.550000	4.910000
22420	4.673496	4.813213

22421	4.730000	4.940000
22422	4.790000	4.630000

	review_scores_communication	review_scores_location \
0	5.000000	4.710000
1	4.900000	4.900000
2	4.800000	4.800000
3	4.808233	4.721844
4	4.808233	4.721844
...	...	...
22418	4.900000	4.650000
22419	4.550000	4.910000
22420	4.800491	4.710323
22421	5.000000	4.850000
22422	4.680000	4.740000

	review_scores_value	reviews_per_month	room_Entire home/apt \
0	4.940000	0.520000	1
1	4.630000	3.810000	0
2	4.200000	2.140000	1
3	4.609505	1.245801	0
4	4.609505	1.245801	0
...	...	...	...
22418	4.710000	3.010000	0
22419	4.450000	0.150000	0
22420	4.604796	1.234094	0
22421	4.760000	0.980000	1
22422	4.680000	0.670000	0

	room_Hotel room	room_Private room	room_Shared room	Air conditioning \
0	0	0	0	0
1	0	1	0	1
2	0	0	0	1
3	0	1	0	0
4	1	0	0	1
...	...	...	...	...
22418	0	1	0	0
22419	0	1	0	1
22420	0	1	0	1
22421	0	0	0	1
22422	0	1	0	0

	Kitchen	Dedicated workspace	Heating	Hot water	Refrigerator \
0	1	0	1	1	1
1	1	0	0	1	0
2	0	1	1	1	0
3	1	1	1	1	1

4	0	1	1	1	0
...	...	...	...	...	...
22418	1	1	1	1	1
22419	0	1	1	0	0
22420	1	1	1	1	1
22421	1	1	0	1	1
22422	0	1	0	1	1

	Free street parking	Self check-in	Shampoo	Washer
0	1	1	1	0
1	1	1	1	0
2	0	1	1	0
3	0	0	0	1
4	0	1	1	0
...	...	...	...	...
22418	1	1	1	0
22419	0	1	0	0
22420	0	0	0	0
22421	0	0	1	0
22422	1	1	1	0

[22423 rows x 51 columns]

### 1.3 Normalize the dataset

```
[ ]: # Initialize MinMaxScaler
scaler = MinMaxScaler()

# Normalize all columns
combined_df = combined_df.astype(float)
combined_df[:] = scaler.fit_transform(combined_df)

combined_df
```

[ ]:	latitude	longitude	host_since	host_response_time \
0	0.448134	0.579717	0.422237	0.333333
1	0.337268	0.531656	0.939973	0.000000
2	0.288168	0.407848	0.870055	0.000000
3	0.422369	0.553485	0.249147	0.000000
4	0.600485	0.495148	0.396487	0.000000
...	...	...	...	...
22418	0.334759	0.558371	0.441337	0.000000
22419	0.786703	0.569332	0.473226	1.000000
22420	0.619740	0.584566	0.249147	0.000000
22421	0.684179	0.496282	0.212995	0.000000
22422	0.475699	0.912347	0.542974	0.000000

	host_response_rate	host_acceptance_rate	host_is_superhost	\
0	1.00	1.00	1.0	
1	1.00	0.98	1.0	
2	1.00	1.00	0.0	
3	0.99	0.23	0.0	
4	0.93	0.95	0.0	
...	...	...	...	
22418	1.00	1.00	0.0	
22419	0.00	0.33	0.0	
22420	0.99	0.23	0.0	
22421	1.00	1.00	1.0	
22422	1.00	1.00	0.0	

	host_listings_count	host_total_listings_count	host_has_profile_pic	\
0	0.000223	0.000111	1.0	
1	0.000000	0.000000	1.0	
2	0.011351	0.005988	1.0	
3	0.161585	0.148037	1.0	
4	0.157133	0.271901	1.0	
...	...	...	...	
22418	0.000223	0.000111	1.0	
22419	0.001335	0.000776	1.0	
22420	0.161585	0.148037	1.0	
22421	0.000000	0.000222	1.0	
22422	0.000445	0.000222	1.0	

	host_identity_verified	calculated_host_listings_count	accommodates	\
0	1.0	0.000000	0.200000	
1	1.0	0.000000	0.066667	
2	1.0	0.058286	0.066667	
3	1.0	0.820571	0.000000	
4	1.0	0.082286	0.066667	
...	...	...	...	
22418	1.0	0.001143	0.066667	
22419	0.0	0.006857	0.000000	
22420	1.0	0.820571	0.000000	
22421	1.0	0.000000	0.066667	
22422	1.0	0.002286	0.066667	

	bathrooms	bedrooms	beds	has_availability	availability_30	\
0	0.173913	0.222222	0.1250	1.0	0.400000	
1	0.086957	0.111111	0.1250	1.0	0.333333	
2	0.086957	0.000000	0.0625	1.0	0.566667	
3	0.130435	0.444444	0.0625	1.0	0.000000	
4	0.086957	0.111111	0.0625	1.0	0.133333	
...	...	...	...	...	...	
22418	0.086957	0.222222	0.3750	1.0	1.000000	

22419	0.086957	0.111111	0.0625	1.0	0.000000
22420	0.260870	0.555556	0.0625	1.0	0.100000
22421	0.086957	0.111111	0.0625	1.0	0.600000
22422	0.086957	0.111111	0.0625	1.0	1.000000

	availability_60	availability_90	availability_365	instant_bookable	\
0	0.700000	0.777778	0.191781	0.0	
1	0.333333	0.544444	0.887671	0.0	
2	0.733333	0.777778	0.400000	1.0	
3	0.000000	0.000000	0.304110	0.0	
4	0.216667	0.244444	0.660274	1.0	
...	...	...	...	...	
22418	1.000000	1.000000	1.000000	0.0	
22419	0.016667	0.011111	0.265753	0.0	
22420	0.550000	0.700000	0.435616	0.0	
22421	0.583333	0.388889	0.095890	0.0	
22422	1.000000	1.000000	0.493151	0.0	

	minimum_nights	maximum_nights	number_of_reviews	\
0	0.058116	0.112411	0.017517	
1	0.000000	0.002800	0.015456	
2	0.000000	0.002800	0.002576	
3	0.058116	0.036404	0.000000	
4	0.000000	0.112411	0.000000	
...	...	...	...	
22418	0.058116	0.112411	0.167439	
22419	0.058116	0.036404	0.005667	
22420	0.058116	0.036404	0.000000	
22421	0.058116	0.002900	0.017002	
22422	0.058116	0.008901	0.009789	

	number_of_reviews_ltm	number_of_reviews_l30d	first_review	\
0	0.002822	0.006803	0.640459	
1	0.016930	0.040816	0.944523	
2	0.002822	0.013605	0.973852	
3	0.000000	0.000000	1.000000	
4	0.000000	0.000000	1.000000	
...	...	...	...	
22418	0.005079	0.000000	0.413428	
22419	0.001129	0.000000	0.598057	
22420	0.000000	0.000000	1.000000	
22421	0.003386	0.000000	0.807951	
22422	0.001129	0.000000	0.836042	

	last_review	review_scores_rating	review_scores_accuracy	\
0	0.978017	1.000000	1.000000	
1	0.982879	0.957500	0.967500	



2	0.979497	0.900000	0.950000
3	1.000000	0.929848	0.935703
4	1.000000	0.929848	0.935703
...	...	...	...
22418	0.921792	0.927500	0.937500
22419	0.963221	0.840000	0.887500
22420	1.000000	0.927862	0.937101
22421	0.969774	0.970000	0.985000
22422	0.950116	0.935000	0.920000

	review_scores_cleanliness	review_scores_checkin \
0	0.992500	1.000000
1	0.982500	0.950000
2	0.800000	0.950000
3	0.919910	0.956577
4	0.919910	0.956577
...	...	...
22418	0.910000	0.970000
22419	0.887500	0.977500
22420	0.918374	0.953303
22421	0.932500	0.985000
22422	0.947500	0.907500

	review_scores_communication	review_scores_location \
0	1.000000	0.927500
1	0.975000	0.975000
2	0.950000	0.950000
3	0.952058	0.930461
4	0.952058	0.930461
...	...	...
22418	0.975000	0.912500
22419	0.887500	0.977500
22420	0.950123	0.927581
22421	1.000000	0.962500
22422	0.920000	0.935000

	review_scores_value	reviews_per_month	room_Entire home/apt \
0	0.985000	0.004633	1.0
1	0.907500	0.034517	0.0
2	0.800000	0.019348	1.0
3	0.902376	0.011225	0.0
4	0.902376	0.011225	0.0
...	...	...	...
22418	0.927500	0.027250	0.0
22419	0.862500	0.001272	0.0
22420	0.901199	0.011119	0.0
22421	0.940000	0.008811	1.0

22422                    0.920000                    0.005995                    0.0

	room_Hotel	room	room_Private	room	room_Shared	room	Air conditioning \
0		0.0		0.0		0.0	0.0
1		0.0		1.0		0.0	1.0
2		0.0		0.0		0.0	1.0
3		0.0		1.0		0.0	0.0
4		1.0		0.0		0.0	1.0
...		...		...		...	
22418		0.0		1.0		0.0	0.0
22419		0.0		1.0		0.0	1.0
22420		0.0		1.0		0.0	1.0
22421		0.0		0.0		0.0	1.0
22422		0.0		1.0		0.0	0.0

	Kitchen	Dedicated workspace	Heating	Hot water	Refrigerator \
0	1.0	0.0	1.0	1.0	1.0
1	1.0	0.0	0.0	1.0	0.0
2	0.0	1.0	1.0	1.0	0.0
3	1.0	1.0	1.0	1.0	1.0
4	0.0	1.0	1.0	1.0	0.0
...	...	...	...	...	...
22418	1.0	1.0	1.0	1.0	1.0
22419	0.0	1.0	1.0	0.0	0.0
22420	1.0	1.0	1.0	1.0	1.0
22421	1.0	1.0	0.0	1.0	1.0
22422	0.0	1.0	0.0	1.0	1.0

	Free street parking	Self check-in	Shampoo	Washer
0	1.0	1.0	1.0	0.0
1	1.0	1.0	1.0	0.0
2	0.0	1.0	1.0	0.0
3	0.0	0.0	0.0	1.0
4	0.0	1.0	1.0	0.0
...	...	...	...	...
22418	1.0	1.0	1.0	0.0
22419	0.0	1.0	0.0	0.0
22420	0.0	0.0	0.0	0.0
22421	0.0	0.0	1.0	0.0
22422	1.0	1.0	1.0	0.0

[22423 rows x 51 columns]

## 2 Unsupervised Learning: Neighborhood Analysis

### 2.1 Use KNN to find averaged neighborhood representations

```
[ ]: # Find k1 nearest neighbors
# Compute average representations of each property according to its k1 nearest
↪ neighbors
def knn(combined_df, k1):
    # Extract longitude and latitude for KNN (ignore them when averaging
    ↪ features)
    coordinates = combined_df[['longitude', 'latitude']].values

    # Initialize the KNN model (for 500 neighbors)
    knn = NearestNeighbors(n_neighbors=k1)

    # Fit the KNN model
    knn.fit(coordinates)

    # Find the 20 nearest neighbors for each property (including itself)
    distances, indices = knn.kneighbors(coordinates)

    # Identify the feature columns
    feature_columns = combined_df.columns[1:] # All columns starting from
    ↪ index 2 onward

    # Create an empty array to store the averaged features
    averaged_features = np.zeros((combined_df.shape[0], len(feature_columns)))

    # For each property, average its own features and those of the 200 nearest
    ↪ neighbors
    for i in range(combined_df.shape[0]):
        # Get the indices of the neighbors (including the property itself)
        neighbor_indices = indices[i]

        # Extract the features for the neighbors
        neighbor_features = combined_df.iloc[neighbor_indices][feature_columns]

        # Ensure the neighbor features are in the same order as feature_columns
        neighbor_features = neighbor_features[feature_columns]

        # Average the features of the property and its neighbors, handling
        ↪ missing values
        averaged_features[i] = neighbor_features.mean(axis=0, skipna=True)

    # Create a DataFrame for the averaged features
    averaged_df = pd.DataFrame(averaged_features, columns=feature_columns)
```

```

    return averaged_df

# Use k1 = 500 (found via experimentation to produce well defined clusters
↳ (little overlap))
averaged_df = knn(combined_df, 500)

```

## 2.2 K-means clustering on averaged neighbor representations

```

[ ]: # K-means on averaged neighborhood representations, computes k2 clusters and
↳ returns the new combined dataframe with

def k_means(k2, averaged_df):

    # Step 1: Apply KMeans clustering on the averaged features
    kmeans = KMeans(n_clusters=k2, random_state=42) # 10 clusters
    averaged_df['cluster'] = kmeans.fit_predict(averaged_df) # Assign clusters
↳ to averaged_df

    # Map the clusters back to the original dataframe (df)
    df = combined_df.merge(averaged_df[['cluster']], left_index=True,
↳ right_index=True)

    # Step 3: Retrieve cluster centroids
    centroids = kmeans.cluster_centers_

    return df, centroids

def plot_cluster(df, k2):
    # Plot the clusters and their centroids
    plt.figure(figsize=(10, 6))

    # Define a colormap with more distinct colors
    cmap = plt.cm.get_cmap('tab20', 20) # Use the 'tab20' colormap with 20
↳ distinct colors

    # Plot each property with its assigned cluster (color-coded)
    scatter = plt.scatter(df['longitude'], df['latitude'], c=df['cluster'],
↳ cmap=cmap, alpha=0.6, s=50)

    # Add a colorbar for reference
    plt.colorbar(scatter, label='Cluster ID')

    # Add labels and title
    plt.xlabel('Longitude', fontsize=14)
    plt.ylabel('Latitude', fontsize=14)
    plt.title(f'K-Means Clustering of Properties ({k2} Clusters)', fontsize=16)

```

```
plt.legend()

# Display the plot
plt.show()

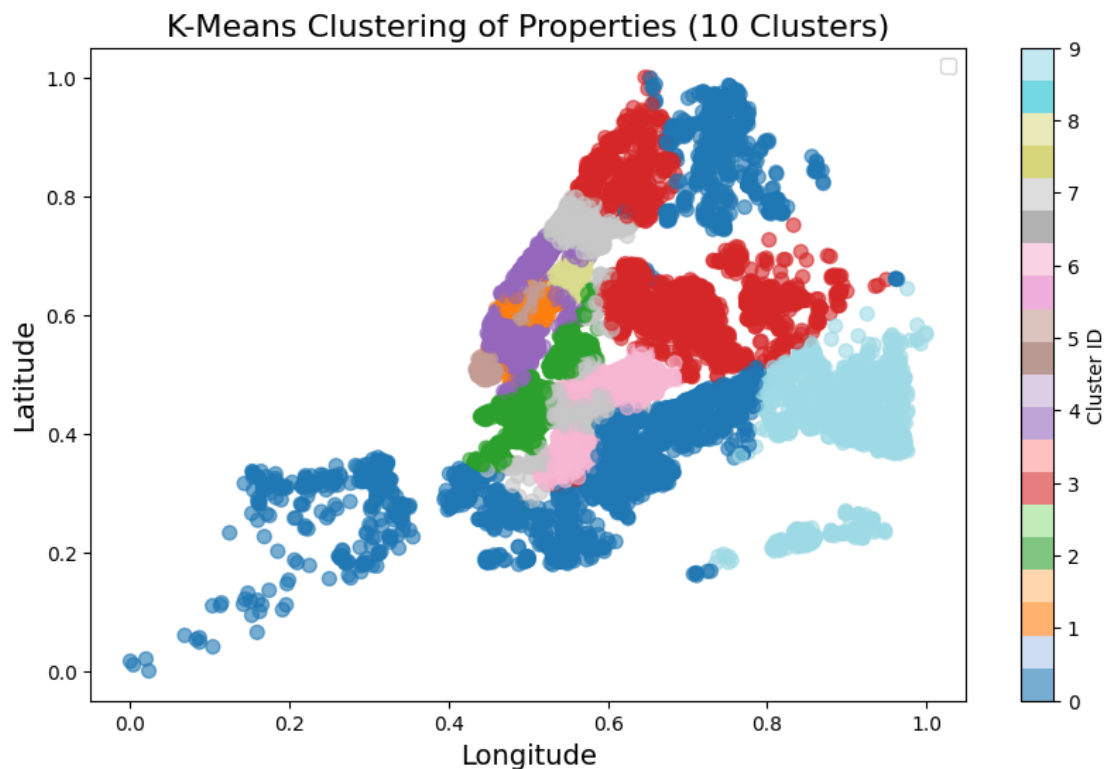
# Use k2 = 10 (found through experimentation)
df, centroids = k_means(10, averaged_df)
plot_cluster(df,10)
```

/var/folders/2f/7sxq51rj0xd8mgyvzzj2tgy80000gn/T/ipykernel\_9234/1503335994.py:22  
: MatplotlibDeprecationWarning: The get\_cmap function was deprecated in  
Matplotlib 3.7 and will be removed in 3.11. Use ``matplotlib.colormaps[name]``  
or ``matplotlib.colormaps.get\_cmap()`` or ``pyplot.get\_cmap()`` instead.

cmap = plt.cm.get\_cmap('tab20', 20) # Use the 'tab20' colormap with 20  
distinct colors

/var/folders/2f/7sxq51rj0xd8mgyvzzj2tgy80000gn/T/ipykernel\_9234/1503335994.py:34  
: UserWarning: No artists with labels found to put in legend. Note that artists  
whose label start with an underscore are ignored when legend() is called with no  
argument.

```
plt.legend()
```



```
[ ]: def train_test_split(df, train_size):

    # Add 'price' and 'id' columns back after manipulation
    df['price'] = price_column
    df['id'] = id_column

    # Split the combined_df back into train_df and test_df
    train_df = df.iloc[:train_size].reset_index(drop=True)
    test_df = df.iloc[train_size:].reset_index(drop=True)

    train_df = train_df.drop(columns='id').reset_index(drop=True)
    test_df = test_df.drop(columns='price').reset_index(drop=True)

    return train_df, test_df

# Example

train_df, test_df = train_test_split(df, train_size)

train_df
```

```
[ ]:
```

	latitude	longitude	host_since	host_response_time \
0	0.448134	0.579717	0.422237	0.333333
1	0.337268	0.531656	0.939973	0.000000
2	0.288168	0.407848	0.870055	0.000000
3	0.422369	0.553485	0.249147	0.000000
4	0.600485	0.495148	0.396487	0.000000
...	...	...	...	...
15691	0.497321	0.456068	0.812415	0.000000
15692	0.927255	0.634970	0.619372	0.000000
15693	0.644912	0.511651	0.595839	0.000000
15694	0.572396	0.457341	0.519952	0.000000
15695	0.310940	0.320027	0.742838	0.000000

	host_response_rate	host_acceptance_rate	host_is_superhost \
0	1.00	1.00	1.0
1	1.00	0.98	1.0
2	1.00	1.00	0.0
3	0.99	0.23	0.0
4	0.93	0.95	0.0
...	...	...	...
15691	0.99	0.99	1.0
15692	1.00	0.67	0.0
15693	1.00	0.98	1.0
15694	1.00	0.96	0.0
15695	1.00	0.83	1.0

	host_listings_count	host_total_listings_count	host_has_profile_pic	\
0	0.000223	0.000111	1.0	
1	0.000000	0.000000	1.0	
2	0.011351	0.005988	1.0	
3	0.161585	0.148037	1.0	
4	0.157133	0.271901	1.0	
...	...	...	...	
15691	0.003116	0.001552	1.0	
15692	0.000223	0.000111	1.0	
15693	0.006009	0.003770	1.0	
15694	1.000000	0.530384	1.0	
15695	0.001113	0.000887	1.0	

	host_identity_verified	calculated_host_listings_count	accommodates	\
0	1.0	0.000000	0.200000	
1	1.0	0.000000	0.066667	
2	1.0	0.058286	0.066667	
3	1.0	0.820571	0.000000	
4	1.0	0.082286	0.066667	
...	...	...	...	
15691	1.0	0.016000	0.200000	
15692	1.0	0.001143	0.066667	
15693	1.0	0.017143	0.133333	
15694	1.0	1.000000	0.066667	
15695	1.0	0.005714	0.000000	

	bathrooms	bedrooms	beds	has_availability	availability_30	\
0	0.173913	0.222222	0.1250	1.0	0.400000	
1	0.086957	0.111111	0.1250	1.0	0.333333	
2	0.086957	0.000000	0.0625	1.0	0.566667	
3	0.130435	0.444444	0.0625	1.0	0.000000	
4	0.086957	0.111111	0.0625	1.0	0.133333	
...	...	...	...	...	...	
15691	0.086957	0.111111	0.2500	1.0	0.200000	
15692	0.086957	0.111111	0.0625	1.0	0.000000	
15693	0.086957	0.111111	0.0625	1.0	0.800000	
15694	0.086957	0.111111	0.0625	1.0	0.000000	
15695	0.086957	0.111111	0.0625	1.0	0.966667	

	availability_60	availability_90	availability_365	instant_bookable	\
0	0.700000	0.777778	0.191781	0.0	
1	0.333333	0.544444	0.887671	0.0	
2	0.733333	0.777778	0.400000	1.0	
3	0.000000	0.000000	0.304110	0.0	
4	0.216667	0.244444	0.660274	1.0	
...	...	...	...	...	
15691	0.200000	0.244444	0.380822	1.0	

15692	0.000000	0.000000	0.591781	0.0
15693	0.450000	0.633333	0.906849	1.0
15694	0.000000	0.000000	0.676712	0.0
15695	0.983333	0.988889	0.490411	0.0

	minimum_nights	maximum_nights	number_of_reviews	\
0	0.058116	0.112411	0.017517	
1	0.000000	0.002800	0.015456	
2	0.000000	0.002800	0.002576	
3	0.058116	0.036404	0.000000	
4	0.000000	0.112411	0.000000	
...	...	...	...	
15691	0.000000	0.036404	0.016486	
15692	0.058116	0.112411	0.003091	
15693	0.000000	0.036404	0.000515	
15694	0.060120	0.112411	0.000000	
15695	0.058116	0.015902	0.010819	

	number_of_reviews_ltm	number_of_reviews_l30d	first_review	\
0	0.002822	0.006803	0.640459	
1	0.016930	0.040816	0.944523	
2	0.002822	0.013605	0.973852	
3	0.000000	0.000000	1.000000	
4	0.000000	0.000000	1.000000	
...	...	...	...	
15691	0.007336	0.006803	0.880389	
15692	0.001129	0.000000	0.896466	
15693	0.000000	0.000000	0.918905	
15694	0.000000	0.000000	1.000000	
15695	0.000000	0.000000	0.737809	

	last_review	review_scores_rating	review_scores_accuracy	\
0	0.978017	1.000000	1.000000	
1	0.982879	0.957500	0.967500	
2	0.979497	0.900000	0.950000	
3	1.000000	0.929848	0.935703	
4	1.000000	0.929848	0.935703	
...	...	...	...	
15691	0.978863	0.985000	0.985000	
15692	0.921792	0.832500	0.832500	
15693	0.902980	1.000000	1.000000	
15694	1.000000	0.929848	0.935703	
15695	0.896428	0.952500	0.975000	

	review_scores_cleanliness	review_scores_checkin	\
0	0.99250	1.000000	
1	0.98250	0.950000	



2	0.80000	0.950000
3	0.91991	0.956577
4	0.91991	0.956577
...	...	...
15691	0.97750	0.992500
15692	0.79250	0.792500
15693	1.00000	1.000000
15694	0.91991	0.956577
15695	0.98750	0.965000

	review_scores_communication	review_scores_location \
0	1.000000	0.927500
1	0.975000	0.975000
2	0.950000	0.950000
3	0.952058	0.930461
4	0.952058	0.930461
...	...	...
15691	0.952500	1.000000
15692	0.832500	0.750000
15693	1.000000	1.000000
15694	0.952058	0.930461
15695	0.965000	0.905000

	review_scores_value	reviews_per_month	room_Entire home/apt \
0	0.985000	0.004633	1.0
1	0.907500	0.034517	0.0
2	0.800000	0.019348	1.0
3	0.902376	0.011225	0.0
4	0.902376	0.011225	0.0
...	...	...	...
15691	0.930000	0.014443	1.0
15692	0.832500	0.003088	0.0
15693	1.000000	0.000636	0.0
15694	0.902376	0.011225	1.0
15695	0.952500	0.003997	0.0

	room_Hotel	room_Private	room_Shared	Air conditioning \
0	0.0	0.0	0.0	0.0
1	0.0	1.0	0.0	1.0
2	0.0	0.0	0.0	1.0
3	0.0	1.0	0.0	0.0
4	1.0	0.0	0.0	1.0
...	...	...	...	...
15691	0.0	0.0	0.0	1.0
15692	0.0	1.0	0.0	0.0
15693	0.0	1.0	0.0	1.0
15694	0.0	0.0	0.0	1.0

15695	0.0	1.0	0.0	1.0
-------	-----	-----	-----	-----

	Kitchen	Dedicated workspace	Heating	Hot water	Refrigerator	\
0	1.0	0.0	1.0	1.0	1.0	
1	1.0	0.0	0.0	1.0	0.0	
2	0.0	1.0	1.0	1.0	0.0	
3	1.0	1.0	1.0	1.0	1.0	
4	0.0	1.0	1.0	1.0	0.0	
...	...	...	...	...	...	
15691	0.0	1.0	1.0	1.0	1.0	
15692	1.0	1.0	0.0	1.0	1.0	
15693	1.0	0.0	1.0	0.0	0.0	
15694	1.0	0.0	1.0	1.0	1.0	
15695	1.0	1.0	1.0	1.0	0.0	

	Free street parking	Self check-in	Shampoo	Washer	cluster	price
0	1.0	1.0	1.0	0.0	7	4.0
1	1.0	1.0	1.0	0.0	6	3.0
2	0.0	1.0	1.0	0.0	0	3.0
3	0.0	0.0	0.0	1.0	7	0.0
4	0.0	1.0	1.0	0.0	5	2.0
...	...	...	...	...	...	
15691	0.0	1.0	1.0	0.0	5	5.0
15692	1.0	1.0	1.0	0.0	3	0.0
15693	0.0	1.0	1.0	0.0	5	5.0
15694	0.0	1.0	1.0	0.0	4	5.0
15695	0.0	1.0	0.0	0.0	0	1.0

[15696 rows x 53 columns]

## 2.3 Train individual GOSDT classifiers on each cluster

### 2.3.1 Use grid-search cross validation within each cluster

```
[ ]: target_column = 'price'
print('here')

# Initialize a list to store models and results
models = []

param_grid = {
    'regularization': [0.01, 0.03, 0.05],
    'depth_budget': [3],
    'balance': [True, False],
    'no_repeat_splits': [False]
}
```

```

# Define the hyperparameter grid for GOSDT
# param_grid = {
#     'regularization': [0.01, 0.03, 0.05],
#     'depth_budget': [3, 5, 7],
#     'balance': [True, False],
#     'no_repeat_splits': [True, False]
# }

# Loop over each cluster
for cluster_id in train_df['cluster'].unique():
    cluster_data = train_df[train_df['cluster'] == cluster_id] # Filter data
    ↪for the cluster

    # Separate features and target
    X = cluster_data.drop(columns=['cluster', target_column]) # Drop 'cluster'
    ↪and 'target_column'
    y = cluster_data[target_column] # The target column

    # Perform hyperparameter tuning with 5-fold cross-validation
    best_rmse = float('inf')
    best_model = None
    hyperparam_results = [] # To store results for graphing

    for params in ParameterGrid(param_grid):
        print('here2')
        configuration = {
            'regularization': params['regularization'],
            'depth_budget': params['depth_budget'],
            'balance': params['balance'],
            'no_repeat_splits': params['no_repeat_splits']
        }

        # 5-Fold Cross-Validation
        kf = KFold(n_splits=5, shuffle=True, random_state=42)
        fold_rmse = []

        for train_idx, val_idx in kf.split(X):
            X_train, X_val = X.iloc[train_idx], X.iloc[val_idx]
            y_train, y_val = y.iloc[train_idx], y.iloc[val_idx]

            # Initialize and train the model
            model = gosdt.GOSDT(configuration)
            model.fit(X_train, y_train)

            # Evaluate on validation set

```

```

        predictions = model.predict(X_val)
        rmse = math.sqrt(mean_squared_error(y_val, predictions))
        fold_rmses.append(rmse)

    # Average RMSE across folds
    avg_rmse = np.mean(fold_rmses)
    hyperparam_results.append((params, avg_rmse))

    # Check for the best model
    if avg_rmse < best_rmse:
        best_rmse = avg_rmse
        best_model = model

# Store the best model for this cluster
models.append({
    'cluster': cluster_id,
    'model': best_model,
    'best_rmse': best_rmse,
    'best_params': configuration
})

# Plot RMSE vs Hyperparameters
for hyperparameter in param_grid.keys():
    plt.figure(figsize=(10, 6))
    values, scores = zip(*[(params[hyperparameter], rmse) for params, rmse
    ↪ in hyperparam_results])
    plt.scatter(values, scores, label=f'Cluster {cluster_id}', alpha=0.7)
    plt.plot(values, scores, marker='o')
    plt.title(f'Cluster {cluster_id}: RMSE vs {hyperparameter}',
    ↪ fontsize=14, fontweight='bold')
    plt.xlabel(hyperparameter, fontsize=12)
    plt.ylabel('Average RMSE (5-Fold)', fontsize=12)
    plt.grid(True)
    plt.legend()
    plt.show()

# Final list of models
print("Trained models and their RMSEs:")
for model_info in models:
    print(f"Cluster {model_info['cluster']} - Best RMSE:
    ↪ {model_info['best_rmse']}, Params: {model_info['best_params']}")

```

here

here2

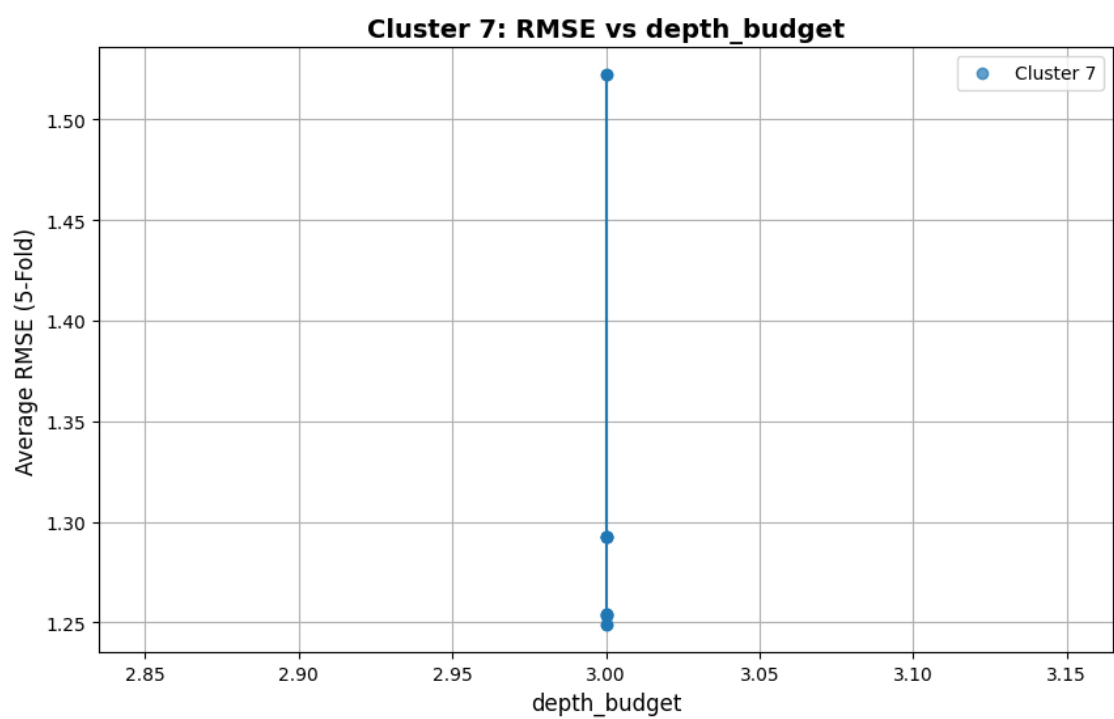
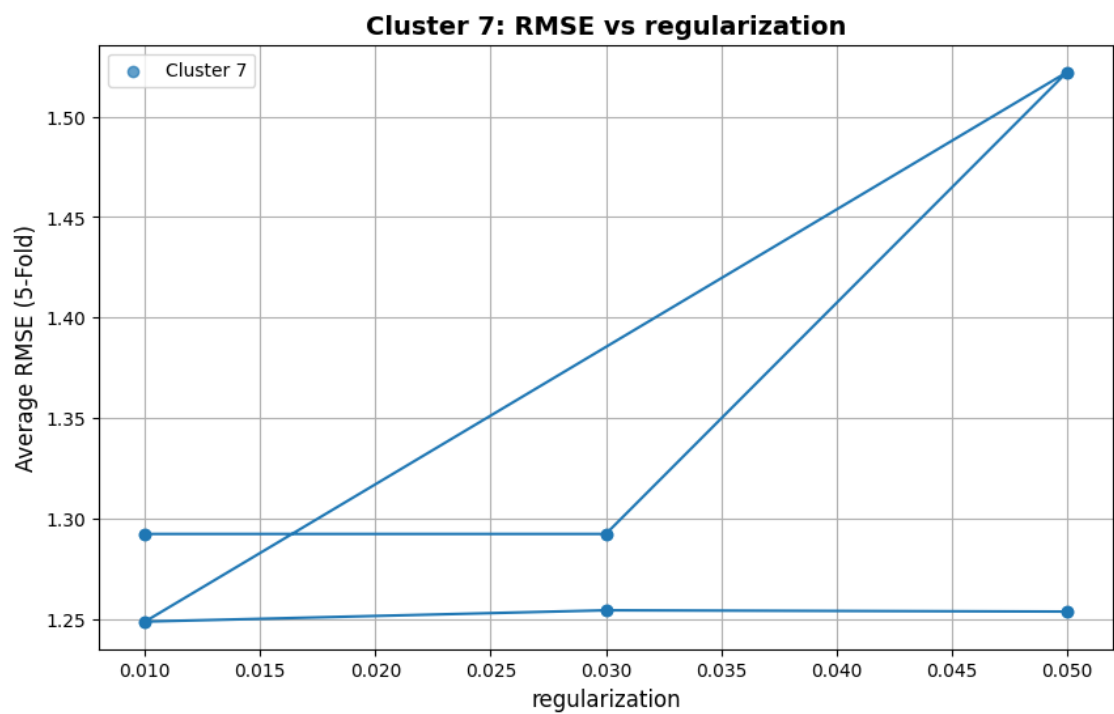
gosdt reported successful execution

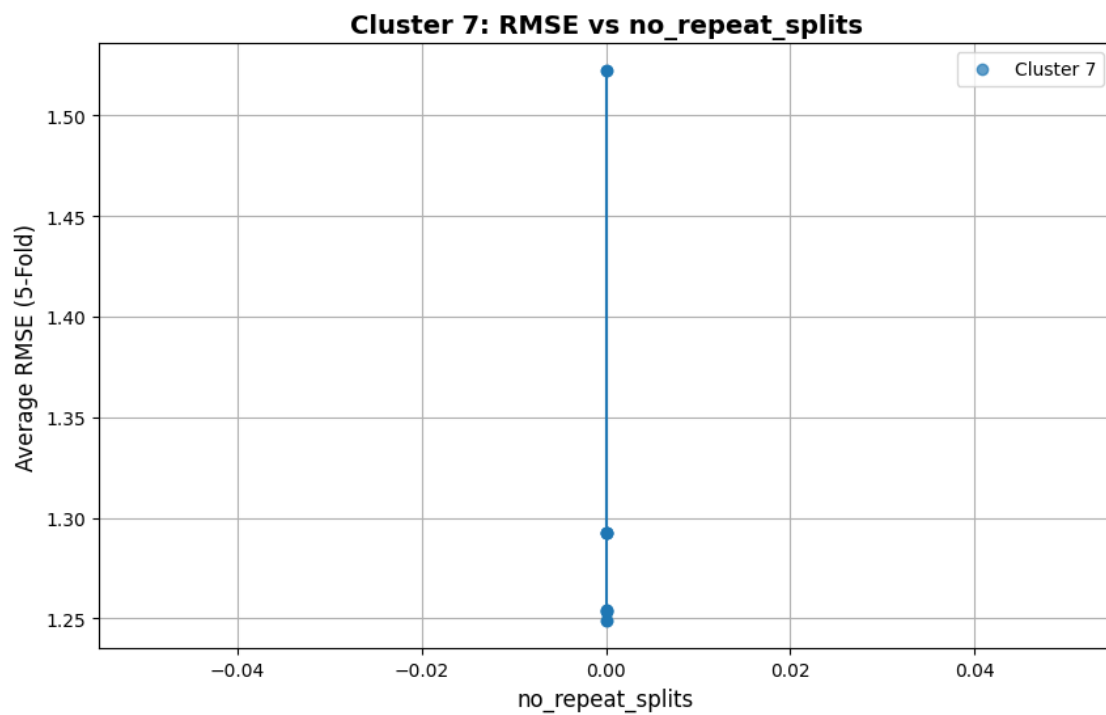
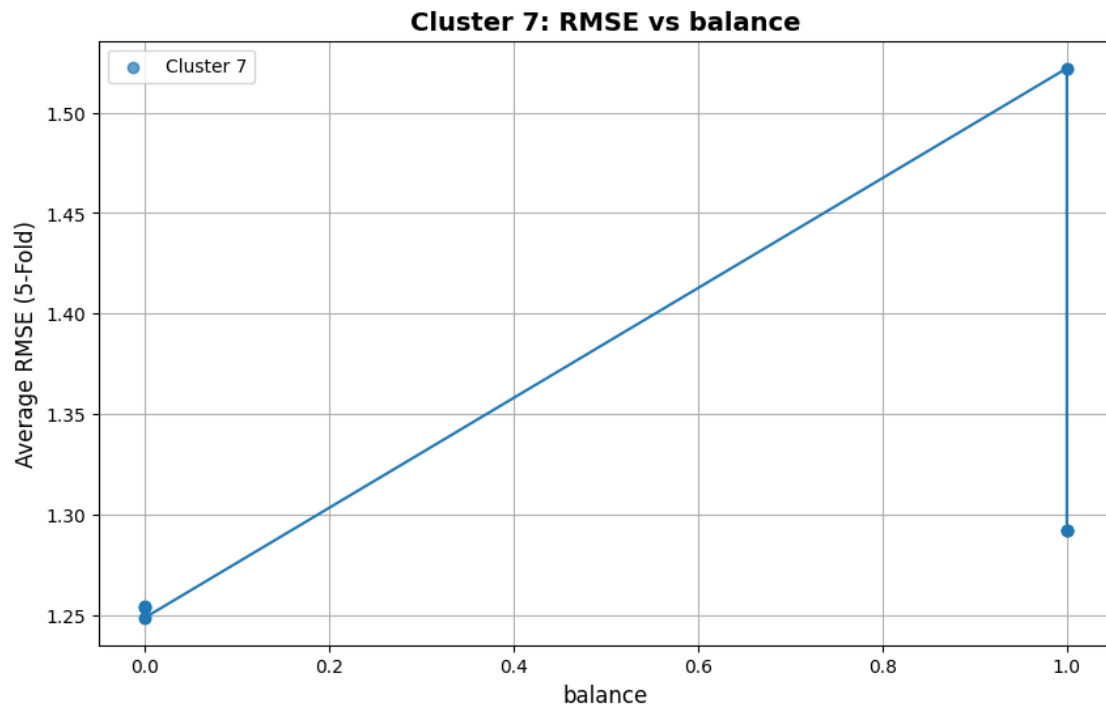
training completed. 0.000/0.000/373.532 (user, system, wall), mem=0 MB

bounds: [0.624494..0.624494] (0.000000) loss=0.584494, iterations=16919

gosdt reported successful execution  
 training completed. 0.000/0.000/362.346 (user, system, wall), mem=0 MB  
 bounds: [0.634125..0.634125] (0.000000) loss=0.594125, iterations=16933  
 gosdt reported successful execution  
 training completed. 0.000/0.000/358.729 (user, system, wall), mem=0 MB  
 bounds: [0.631258..0.631258] (0.000000) loss=0.591258, iterations=16889  
 gosdt reported successful execution  
 training completed. 0.000/0.000/361.688 (user, system, wall), mem=0 MB  
 bounds: [0.624639..0.624639] (0.000000) loss=0.584639, iterations=16841  
 gosdt reported successful execution  
 training completed. 0.000/0.000/354.271 (user, system, wall), mem=0 MB  
 bounds: [0.630765..0.630765] (0.000000) loss=0.590765, iterations=17023  
 here2  
 gosdt reported successful execution  
 training completed. 0.000/0.000/332.019 (user, system, wall), mem=0 MB  
 bounds: [0.704494..0.704494] (0.000000) loss=0.584494, iterations=16919  
 gosdt reported successful execution  
 training completed. 0.000/0.000/341.449 (user, system, wall), mem=0 MB  
 bounds: [0.714125..0.714125] (0.000000) loss=0.594125, iterations=16933  
 gosdt reported successful execution  
 training completed. 0.000/0.000/337.698 (user, system, wall), mem=0 MB  
 bounds: [0.711258..0.711258] (0.000000) loss=0.591258, iterations=16889  
 gosdt reported successful execution  
 training completed. 0.000/0.000/336.077 (user, system, wall), mem=0 MB  
 bounds: [0.704639..0.704639] (0.000000) loss=0.584639, iterations=16841  
 gosdt reported successful execution  
 training completed. 0.000/0.000/366.082 (user, system, wall), mem=0 MB  
 bounds: [0.710765..0.710765] (0.000000) loss=0.590765, iterations=17023  
 here2  
 gosdt reported successful execution  
 training completed. 0.000/0.000/316.584 (user, system, wall), mem=0 MB  
 bounds: [0.776044..0.776044] (0.000000) loss=0.626044, iterations=16919  
 gosdt reported successful execution  
 training completed. 0.000/0.000/328.928 (user, system, wall), mem=0 MB  
 bounds: [0.780894..0.780894] (0.000000) loss=0.630894, iterations=16933  
 gosdt reported successful execution  
 training completed. 0.000/0.000/325.208 (user, system, wall), mem=0 MB  
 bounds: [0.786502..0.786502] (0.000000) loss=0.636502, iterations=16889  
 gosdt reported successful execution  
 training completed. 0.000/0.000/314.781 (user, system, wall), mem=0 MB  
 bounds: [0.779855..0.779855] (0.000000) loss=0.629855, iterations=16841  
 gosdt reported successful execution  
 training completed. 0.000/0.000/333.653 (user, system, wall), mem=0 MB  
 bounds: [0.782772..0.782772] (0.000000) loss=0.632772, iterations=17023  
 here2  
 gosdt reported successful execution  
 training completed. 0.000/0.000/361.340 (user, system, wall), mem=0 MB  
 bounds: [0.607792..0.607792] (0.000000) loss=0.567792, iterations=16919

gosdt reported successful execution  
 training completed. 0.000/0.000/372.902 (user, system, wall), mem=0 MB  
 bounds: [0.624935..0.624935] (0.000000) loss=0.584935, iterations=16933  
 gosdt reported successful execution  
 training completed. 0.000/0.000/367.161 (user, system, wall), mem=0 MB  
 bounds: [0.615286..0.615286] (0.000000) loss=0.575286, iterations=16889  
 gosdt reported successful execution  
 training completed. 0.000/0.000/369.603 (user, system, wall), mem=0 MB  
 bounds: [0.610093..0.610093] (0.000000) loss=0.570093, iterations=16841  
 gosdt reported successful execution  
 training completed. 0.000/0.000/355.987 (user, system, wall), mem=0 MB  
 bounds: [0.615805..0.615805] (0.000000) loss=0.575805, iterations=17023  
 here2  
 gosdt reported successful execution  
 training completed. 0.000/0.000/331.615 (user, system, wall), mem=0 MB  
 bounds: [0.687792..0.687792] (0.000000) loss=0.567792, iterations=16919  
 gosdt reported successful execution  
 training completed. 0.000/0.000/339.433 (user, system, wall), mem=0 MB  
 bounds: [0.702987..0.702987] (0.000000) loss=0.612987, iterations=16933  
 gosdt reported successful execution  
 training completed. 0.000/0.000/339.208 (user, system, wall), mem=0 MB  
 bounds: [0.692804..0.692804] (0.000000) loss=0.602804, iterations=16889  
 gosdt reported successful execution  
 training completed. 0.000/0.000/329.010 (user, system, wall), mem=0 MB  
 bounds: [0.689169..0.689169] (0.000000) loss=0.599169, iterations=16841  
 gosdt reported successful execution  
 training completed. 0.000/0.000/346.302 (user, system, wall), mem=0 MB  
 bounds: [0.693323..0.693323] (0.000000) loss=0.603323, iterations=17023  
 here2  
 gosdt reported successful execution  
 training completed. 0.000/0.000/328.457 (user, system, wall), mem=0 MB  
 bounds: [0.750000..0.750000] (0.000000) loss=0.600000, iterations=16919  
 gosdt reported successful execution  
 training completed. 0.000/0.000/340.419 (user, system, wall), mem=0 MB  
 bounds: [0.762987..0.762987] (0.000000) loss=0.612987, iterations=16933  
 gosdt reported successful execution  
 training completed. 0.000/0.000/336.851 (user, system, wall), mem=0 MB  
 bounds: [0.752804..0.752804] (0.000000) loss=0.602804, iterations=16889  
 gosdt reported successful execution  
 training completed. 0.000/0.000/322.716 (user, system, wall), mem=0 MB  
 bounds: [0.749169..0.749169] (0.000000) loss=0.599169, iterations=16841  
 gosdt reported successful execution  
 training completed. 0.000/0.000/329.884 (user, system, wall), mem=0 MB  
 bounds: [0.753323..0.753323] (0.000000) loss=0.603323, iterations=17023





here2



```

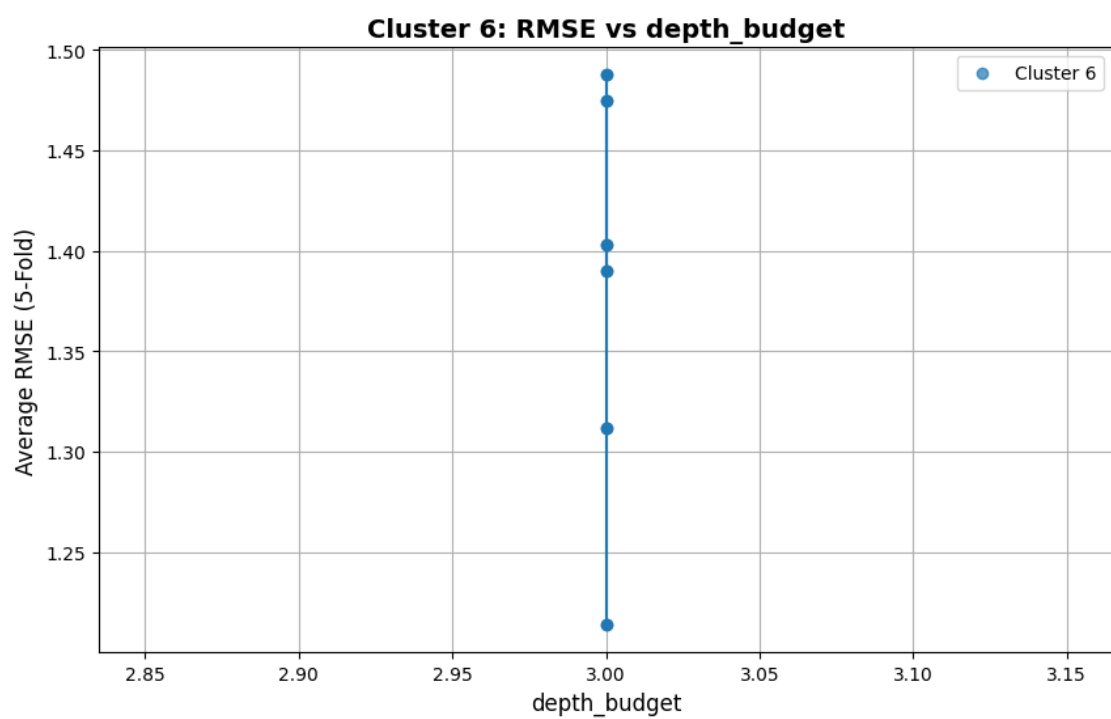
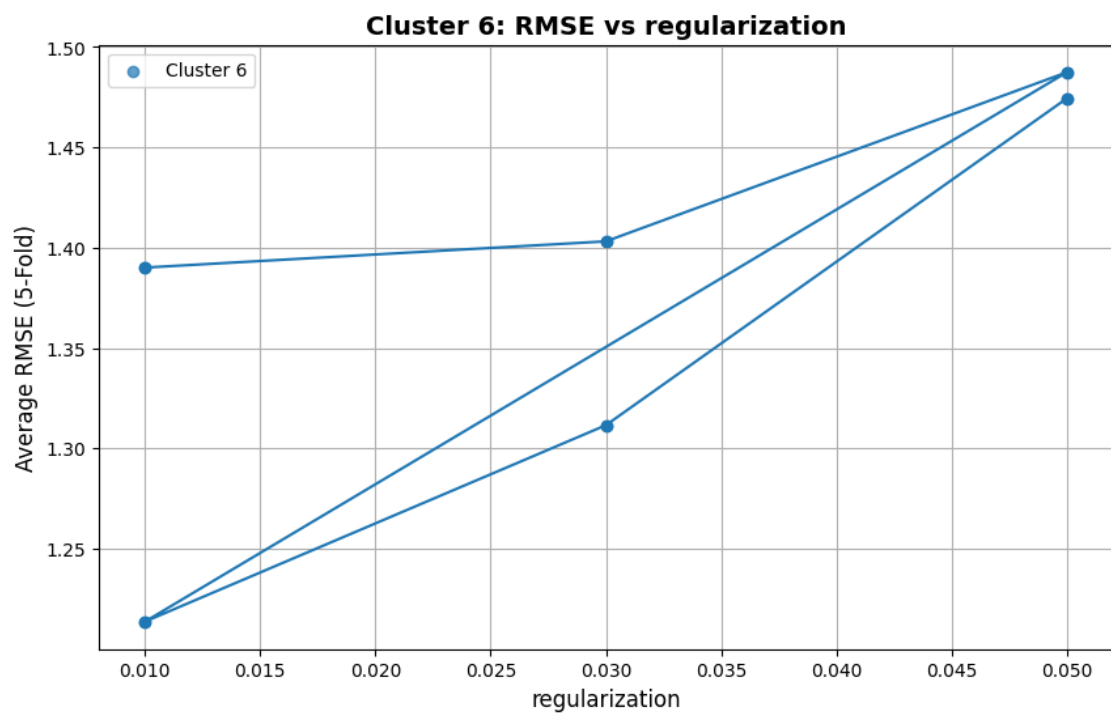
gosdt reported successful execution
training completed. 0.000/0.000/201.836 (user, system, wall), mem=0 MB
bounds: [0.625581..0.625581] (0.000000) loss=0.585581, iterations=13995
gosdt reported successful execution
training completed. 0.000/0.000/198.673 (user, system, wall), mem=0 MB
bounds: [0.630671..0.630671] (0.000000) loss=0.590671, iterations=13869
gosdt reported successful execution
training completed. 0.000/0.000/203.842 (user, system, wall), mem=0 MB
bounds: [0.622915..0.622915] (0.000000) loss=0.582915, iterations=13937
gosdt reported successful execution
training completed. 0.000/0.000/204.824 (user, system, wall), mem=0 MB
bounds: [0.618547..0.618547] (0.000000) loss=0.578547, iterations=13953
gosdt reported successful execution
training completed. 0.000/0.000/198.997 (user, system, wall), mem=0 MB
bounds: [0.620082..0.620082] (0.000000) loss=0.580082, iterations=13897
here2
gosdt reported successful execution
training completed. 0.000/0.000/193.016 (user, system, wall), mem=0 MB
bounds: [0.705581..0.705581] (0.000000) loss=0.585581, iterations=13995
gosdt reported successful execution
training completed. 0.000/0.000/193.047 (user, system, wall), mem=0 MB
bounds: [0.710671..0.710671] (0.000000) loss=0.590671, iterations=13869
gosdt reported successful execution
training completed. 0.000/0.000/194.592 (user, system, wall), mem=0 MB
bounds: [0.702915..0.702915] (0.000000) loss=0.582915, iterations=13937
gosdt reported successful execution
training completed. 0.000/0.000/195.852 (user, system, wall), mem=0 MB
bounds: [0.698060..0.698060] (0.000000) loss=0.608060, iterations=13953
gosdt reported successful execution
training completed. 0.000/0.000/190.423 (user, system, wall), mem=0 MB
bounds: [0.700082..0.700082] (0.000000) loss=0.580082, iterations=13897
here2
gosdt reported successful execution
training completed. 0.000/0.000/184.757 (user, system, wall), mem=0 MB
bounds: [0.768249..0.768249] (0.000000) loss=0.618249, iterations=13995
gosdt reported successful execution
training completed. 0.000/0.000/184.926 (user, system, wall), mem=0 MB
bounds: [0.773523..0.773523] (0.000000) loss=0.623523, iterations=13869
gosdt reported successful execution
training completed. 0.000/0.000/186.969 (user, system, wall), mem=0 MB
bounds: [0.768789..0.768789] (0.000000) loss=0.618788, iterations=13937
gosdt reported successful execution
training completed. 0.000/0.000/187.775 (user, system, wall), mem=0 MB
bounds: [0.758060..0.758060] (0.000000) loss=0.608060, iterations=13953
gosdt reported successful execution
training completed. 0.000/0.000/182.549 (user, system, wall), mem=0 MB
bounds: [0.760273..0.760273] (0.000000) loss=0.610273, iterations=13897
here2

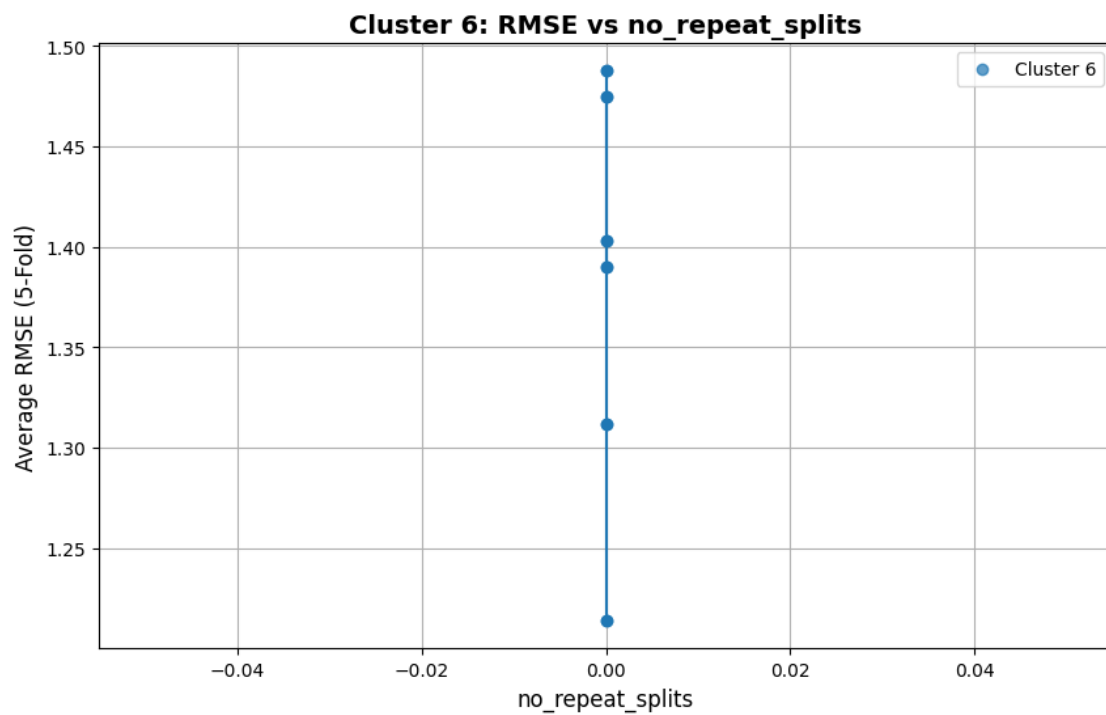
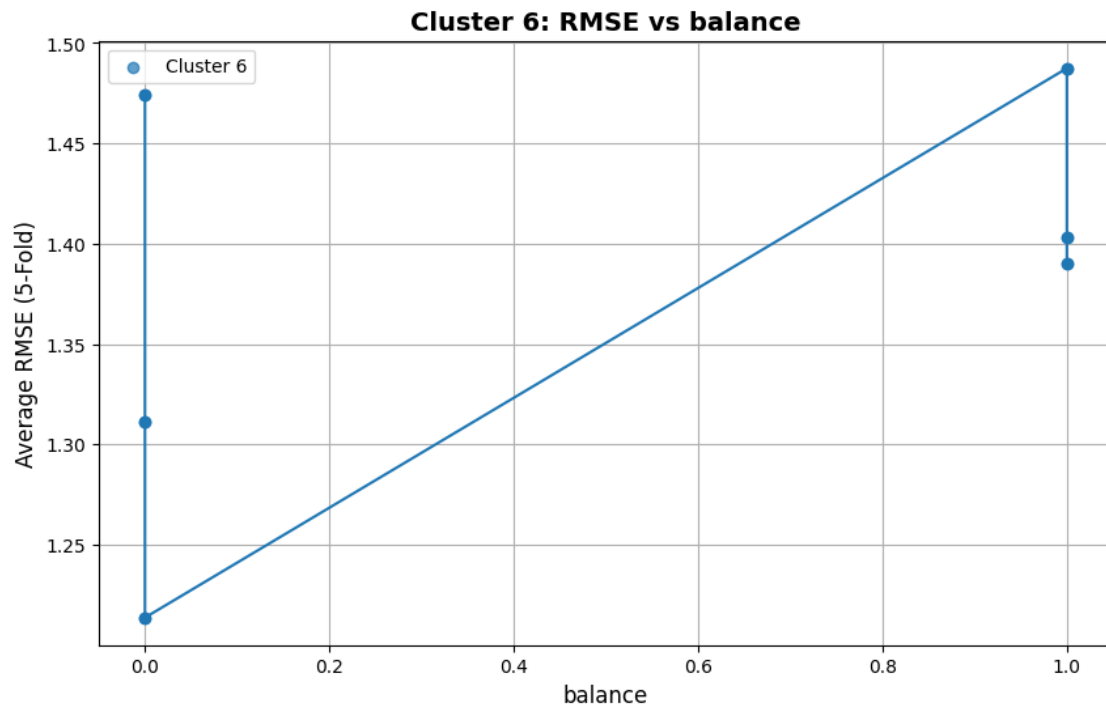
```

```

gosdt reported successful execution
training completed. 0.000/0.000/201.334 (user, system, wall), mem=0 MB
bounds: [0.520104..0.520104] (0.000000) loss=0.480104, iterations=13995
gosdt reported successful execution
training completed. 0.000/0.000/199.667 (user, system, wall), mem=0 MB
bounds: [0.506406..0.506406] (0.000000) loss=0.466406, iterations=13869
gosdt reported successful execution
training completed. 0.000/0.000/202.580 (user, system, wall), mem=0 MB
bounds: [0.523703..0.523703] (0.000000) loss=0.483703, iterations=13937
gosdt reported successful execution
training completed. 0.000/0.000/203.634 (user, system, wall), mem=0 MB
bounds: [0.511317..0.511317] (0.000000) loss=0.471317, iterations=13953
gosdt reported successful execution
training completed. 0.000/0.000/198.823 (user, system, wall), mem=0 MB
bounds: [0.525007..0.525007] (0.000000) loss=0.485007, iterations=13897
here2
gosdt reported successful execution
training completed. 0.000/0.000/192.285 (user, system, wall), mem=0 MB
bounds: [0.591631..0.591631] (0.000000) loss=0.501631, iterations=13995
gosdt reported successful execution
training completed. 0.000/0.000/191.769 (user, system, wall), mem=0 MB
bounds: [0.579237..0.579237] (0.000000) loss=0.489237, iterations=13869
gosdt reported successful execution
training completed. 0.000/0.000/193.135 (user, system, wall), mem=0 MB
bounds: [0.593911..0.593911] (0.000000) loss=0.503911, iterations=13937
gosdt reported successful execution
training completed. 0.000/0.000/193.920 (user, system, wall), mem=0 MB
bounds: [0.586089..0.586089] (0.000000) loss=0.496089, iterations=13953
gosdt reported successful execution
training completed. 0.000/0.000/189.207 (user, system, wall), mem=0 MB
bounds: [0.591943..0.591943] (0.000000) loss=0.531943, iterations=13897
here2
gosdt reported successful execution
training completed. 0.000/0.000/184.964 (user, system, wall), mem=0 MB
bounds: [0.633594..0.633594] (0.000000) loss=0.533594, iterations=13995
gosdt reported successful execution
training completed. 0.000/0.000/184.768 (user, system, wall), mem=0 MB
bounds: [0.622505..0.622505] (0.000000) loss=0.522505, iterations=13869
gosdt reported successful execution
training completed. 0.000/0.000/186.714 (user, system, wall), mem=0 MB
bounds: [0.636506..0.636506] (0.000000) loss=0.536506, iterations=13937
gosdt reported successful execution
training completed. 0.000/0.000/186.766 (user, system, wall), mem=0 MB
bounds: [0.629335..0.629335] (0.000000) loss=0.529335, iterations=13953
gosdt reported successful execution
training completed. 0.000/0.000/181.747 (user, system, wall), mem=0 MB
bounds: [0.631943..0.631943] (0.000000) loss=0.531943, iterations=13897

```





here2

```

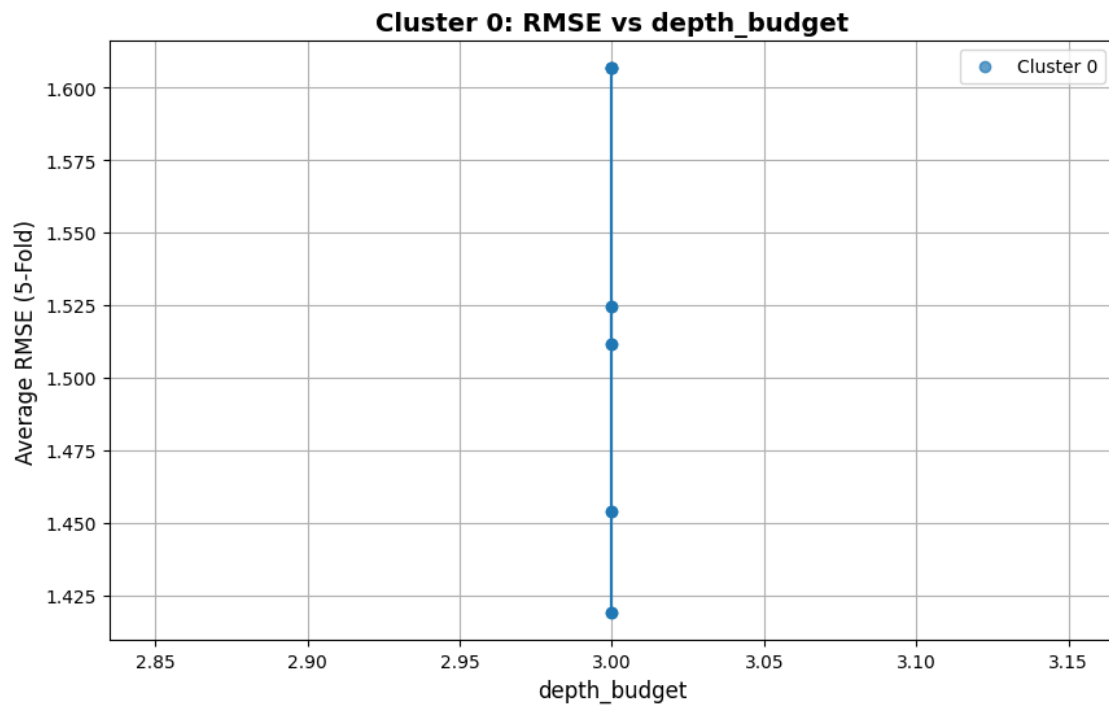
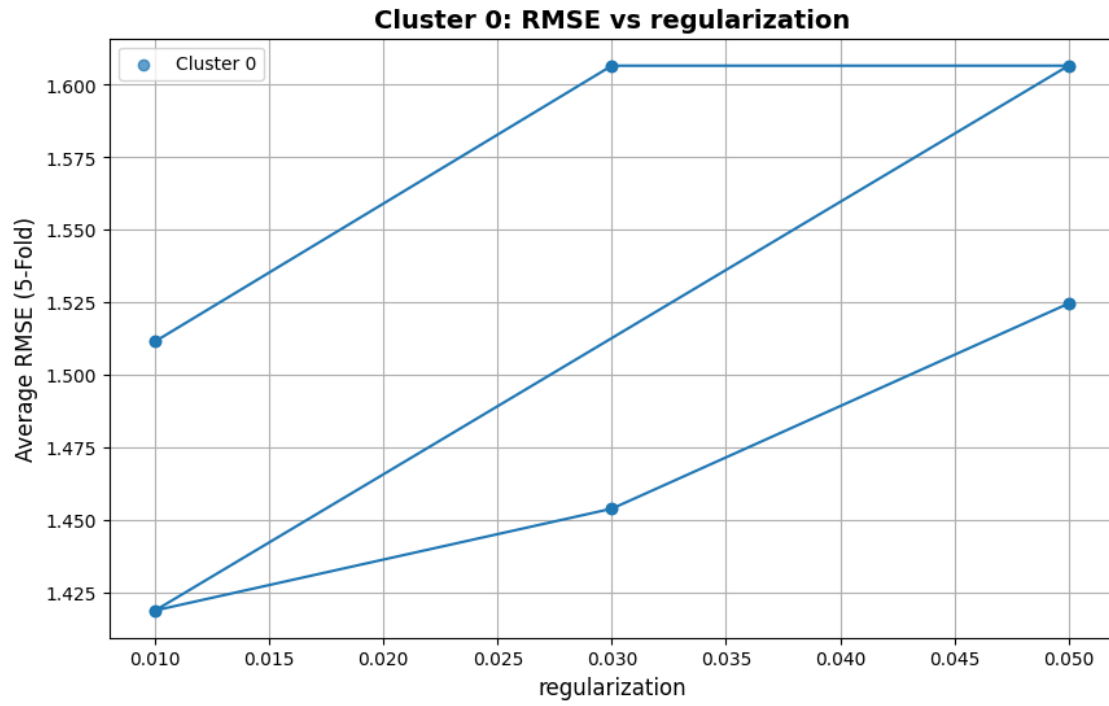
gosdt reported successful execution
training completed. 0.000/0.000/236.320 (user, system, wall), mem=0 MB
bounds: [0.646552..0.646552] (0.000000) loss=0.606552, iterations=15457
gosdt reported successful execution
training completed. 0.000/0.000/239.672 (user, system, wall), mem=0 MB
bounds: [0.650623..0.650623] (0.000000) loss=0.610623, iterations=15427
gosdt reported successful execution
training completed. 0.000/0.000/229.592 (user, system, wall), mem=0 MB
bounds: [0.644750..0.644750] (0.000000) loss=0.604750, iterations=15277
gosdt reported successful execution
training completed. 0.000/0.000/233.638 (user, system, wall), mem=0 MB
bounds: [0.651497..0.651497] (0.000000) loss=0.611497, iterations=15345
gosdt reported successful execution
training completed. 0.000/0.000/240.176 (user, system, wall), mem=0 MB
bounds: [0.651880..0.651880] (0.000000) loss=0.611880, iterations=15309
here2
gosdt reported successful execution
training completed. 0.000/0.000/224.644 (user, system, wall), mem=0 MB
bounds: [0.720419..0.720419] (0.000000) loss=0.630419, iterations=15457
gosdt reported successful execution
training completed. 0.000/0.000/230.246 (user, system, wall), mem=0 MB
bounds: [0.726389..0.726389] (0.000000) loss=0.636389, iterations=15427
gosdt reported successful execution
training completed. 0.000/0.000/226.952 (user, system, wall), mem=0 MB
bounds: [0.716617..0.716617] (0.000000) loss=0.626617, iterations=15277
gosdt reported successful execution
training completed. 0.000/0.000/222.731 (user, system, wall), mem=0 MB
bounds: [0.727004..0.727004] (0.000000) loss=0.637004, iterations=15345
gosdt reported successful execution
training completed. 0.000/0.000/230.139 (user, system, wall), mem=0 MB
bounds: [0.730283..0.730283] (0.000000) loss=0.640283, iterations=15309
here2
gosdt reported successful execution
training completed. 0.000/0.000/217.983 (user, system, wall), mem=0 MB
bounds: [0.780419..0.780419] (0.000000) loss=0.630419, iterations=15457
gosdt reported successful execution
training completed. 0.000/0.000/223.720 (user, system, wall), mem=0 MB
bounds: [0.786389..0.786389] (0.000000) loss=0.636389, iterations=15427
gosdt reported successful execution
training completed. 0.000/0.000/208.737 (user, system, wall), mem=0 MB
bounds: [0.776617..0.776617] (0.000000) loss=0.626617, iterations=15277
gosdt reported successful execution
training completed. 0.000/0.000/213.395 (user, system, wall), mem=0 MB
bounds: [0.787004..0.787004] (0.000000) loss=0.637004, iterations=15345
gosdt reported successful execution
training completed. 0.000/0.000/218.416 (user, system, wall), mem=0 MB
bounds: [0.790283..0.790283] (0.000000) loss=0.640283, iterations=15309
here2

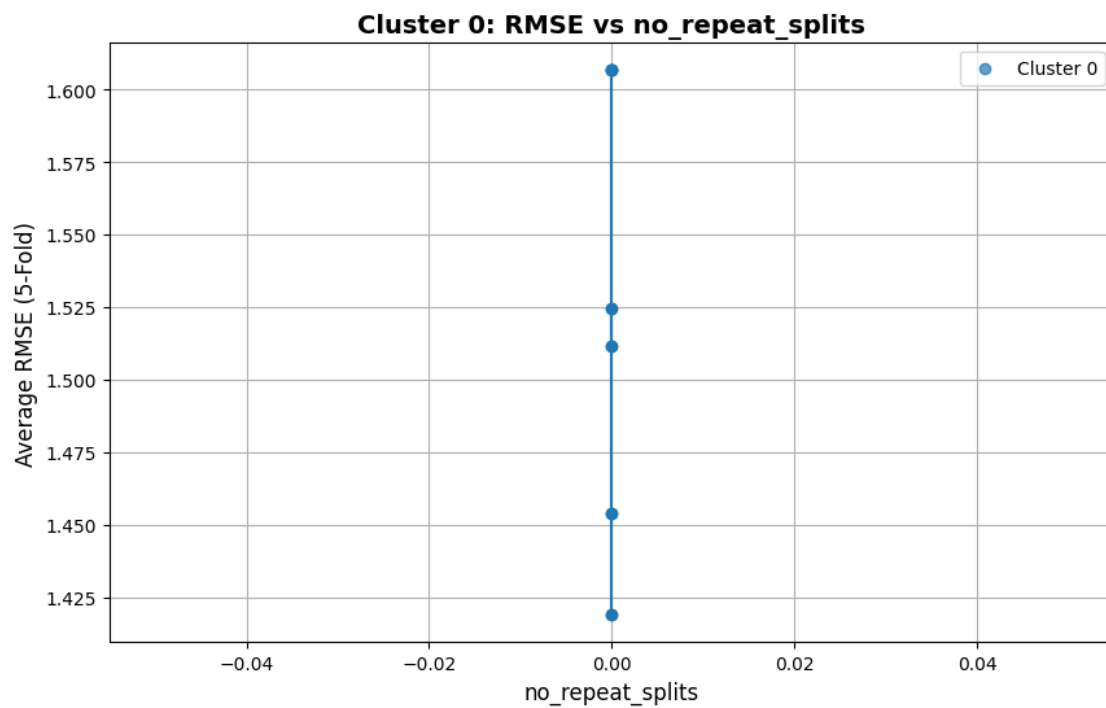
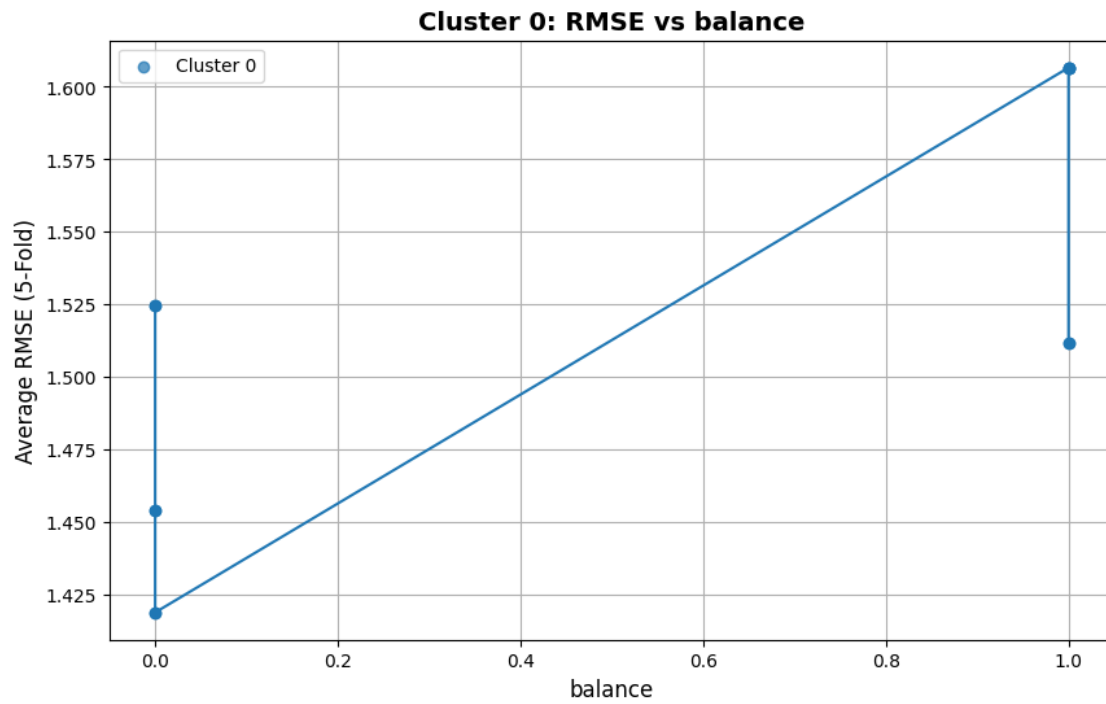
```

```

gosdt reported successful execution
training completed. 0.000/0.000/236.238 (user, system, wall), mem=0 MB
bounds: [0.636071..0.636071] (0.000000) loss=0.596071, iterations=15457
gosdt reported successful execution
training completed. 0.000/0.000/242.402 (user, system, wall), mem=0 MB
bounds: [0.645893..0.645893] (0.000000) loss=0.605893, iterations=15427
gosdt reported successful execution
training completed. 0.000/0.000/233.297 (user, system, wall), mem=0 MB
bounds: [0.634479..0.634479] (0.000000) loss=0.594479, iterations=15277
gosdt reported successful execution
training completed. 0.000/0.000/233.144 (user, system, wall), mem=0 MB
bounds: [0.647975..0.647975] (0.000000) loss=0.607975, iterations=15345
gosdt reported successful execution
training completed. 0.000/0.000/239.425 (user, system, wall), mem=0 MB
bounds: [0.640613..0.640613] (0.000000) loss=0.600613, iterations=15309
here2
gosdt reported successful execution
training completed. 0.000/0.000/225.879 (user, system, wall), mem=0 MB
bounds: [0.705101..0.705101] (0.000000) loss=0.615101, iterations=15457
gosdt reported successful execution
training completed. 0.000/0.000/229.003 (user, system, wall), mem=0 MB
bounds: [0.714923..0.714923] (0.000000) loss=0.624923, iterations=15427
gosdt reported successful execution
training completed. 0.000/0.000/219.259 (user, system, wall), mem=0 MB
bounds: [0.707791..0.707791] (0.000000) loss=0.617791, iterations=15277
gosdt reported successful execution
training completed. 0.000/0.000/223.350 (user, system, wall), mem=0 MB
bounds: [0.718834..0.718834] (0.000000) loss=0.628834, iterations=15345
gosdt reported successful execution
training completed. 0.000/0.000/227.688 (user, system, wall), mem=0 MB
bounds: [0.711472..0.711472] (0.000000) loss=0.621472, iterations=15309
here2
gosdt reported successful execution
training completed. 0.000/0.000/220.261 (user, system, wall), mem=0 MB
bounds: [0.761756..0.761756] (0.000000) loss=0.661756, iterations=15457
gosdt reported successful execution
training completed. 0.000/0.000/220.285 (user, system, wall), mem=0 MB
bounds: [0.774923..0.774923] (0.000000) loss=0.624923, iterations=15427
gosdt reported successful execution
training completed. 0.000/0.000/218.474 (user, system, wall), mem=0 MB
bounds: [0.759509..0.759509] (0.000000) loss=0.659509, iterations=15277
gosdt reported successful execution
training completed. 0.000/0.000/213.071 (user, system, wall), mem=0 MB
bounds: [0.778834..0.778834] (0.000000) loss=0.628834, iterations=15345
gosdt reported successful execution
training completed. 0.000/0.000/215.426 (user, system, wall), mem=0 MB
bounds: [0.763804..0.763804] (0.000000) loss=0.663804, iterations=15309

```



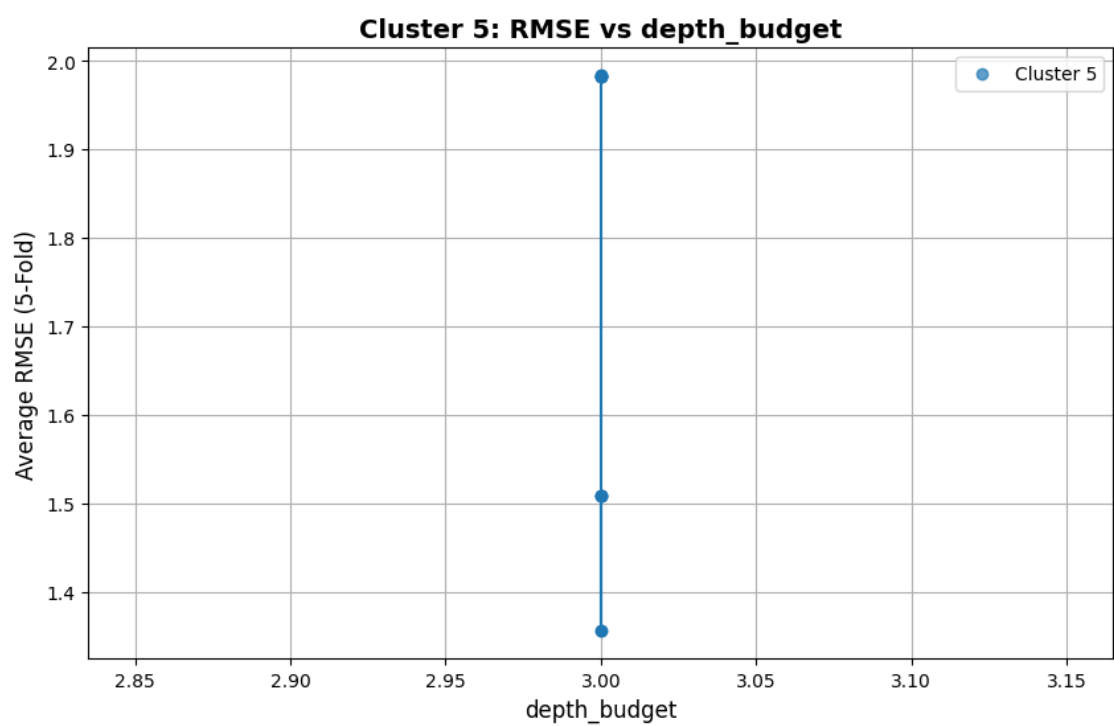
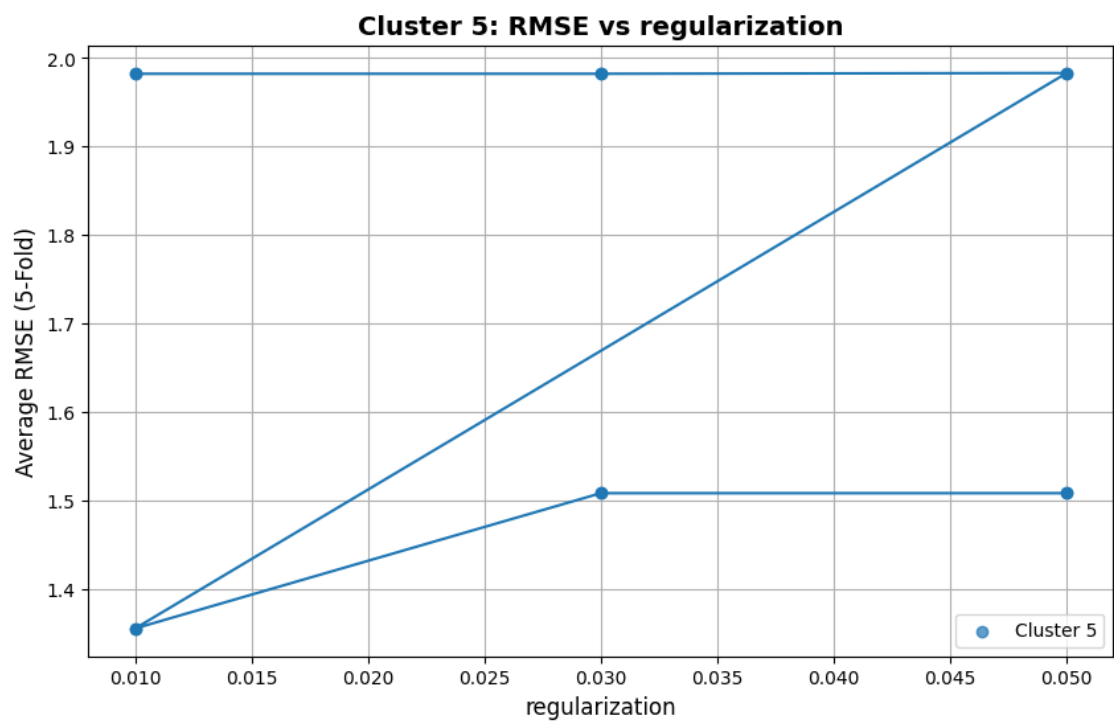


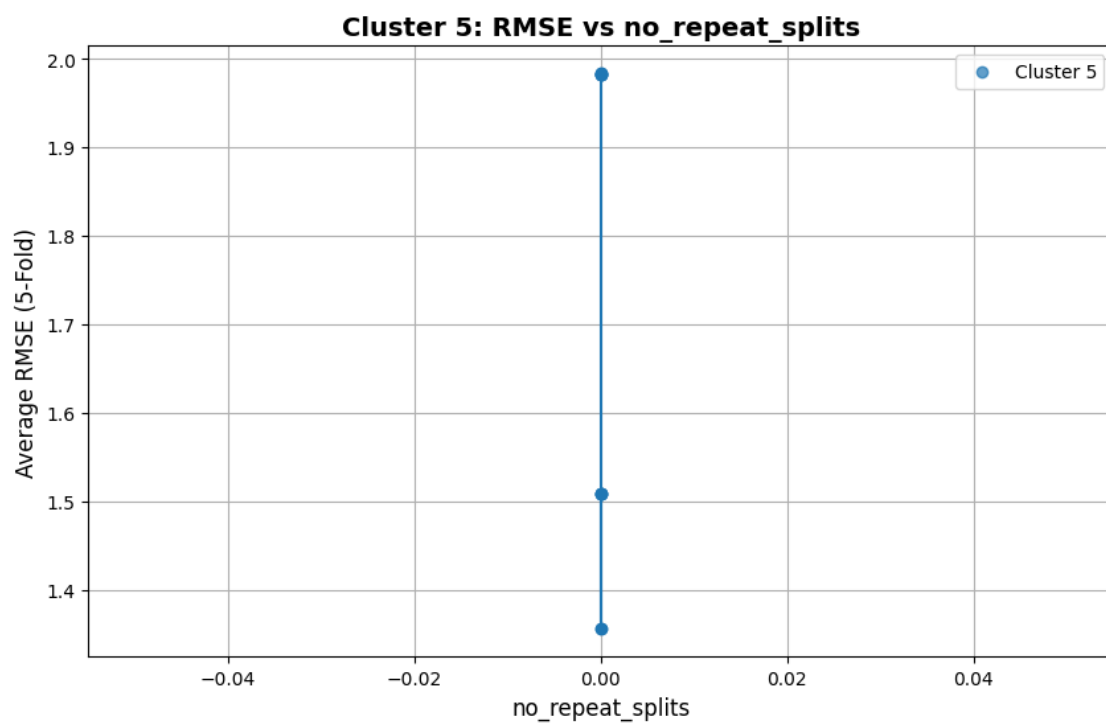
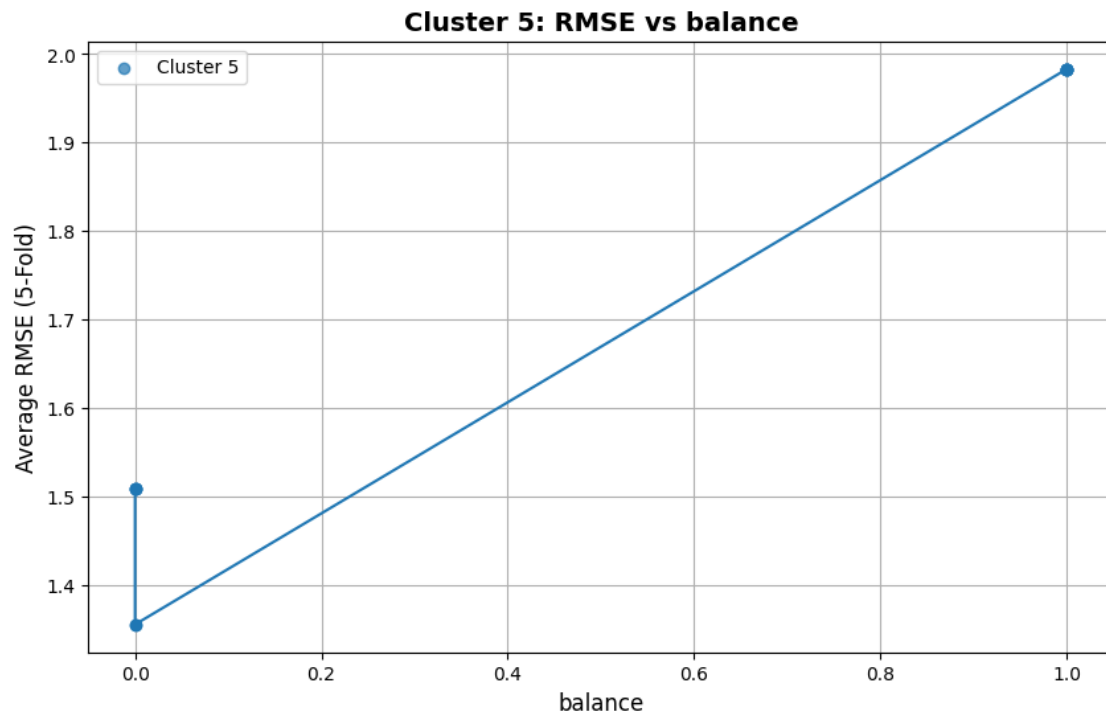
here2  
gosdt reported successful execution



training completed. 0.000/0.000/65.032 (user, system, wall), mem=0 MB  
 bounds: [0.580667..0.580667] (0.000000) loss=0.540667, iterations=8841  
 gosdt reported successful execution  
 training completed. 0.000/0.000/62.823 (user, system, wall), mem=0 MB  
 bounds: [0.593486..0.593486] (0.000000) loss=0.553486, iterations=8715  
 gosdt reported successful execution  
 training completed. 0.000/0.000/66.364 (user, system, wall), mem=0 MB  
 bounds: [0.573052..0.573052] (0.000000) loss=0.533051, iterations=8927  
 gosdt reported successful execution  
 training completed. 0.000/0.000/64.811 (user, system, wall), mem=0 MB  
 bounds: [0.581482..0.581482] (0.000000) loss=0.541482, iterations=8833  
 gosdt reported successful execution  
 training completed. 0.000/0.000/64.096 (user, system, wall), mem=0 MB  
 bounds: [0.580530..0.580530] (0.000000) loss=0.540530, iterations=8793  
 here2  
 gosdt reported successful execution  
 training completed. 0.000/0.000/62.105 (user, system, wall), mem=0 MB  
 bounds: [0.660667..0.660667] (0.000000) loss=0.540667, iterations=8841  
 gosdt reported successful execution  
 training completed. 0.000/0.000/59.845 (user, system, wall), mem=0 MB  
 bounds: [0.673486..0.673486] (0.000000) loss=0.553486, iterations=8715  
 gosdt reported successful execution  
 training completed. 0.000/0.000/63.226 (user, system, wall), mem=0 MB  
 bounds: [0.653051..0.653051] (0.000000) loss=0.533051, iterations=8927  
 gosdt reported successful execution  
 training completed. 0.000/0.000/61.182 (user, system, wall), mem=0 MB  
 bounds: [0.661482..0.661482] (0.000000) loss=0.541482, iterations=8833  
 gosdt reported successful execution  
 training completed. 0.000/0.000/60.635 (user, system, wall), mem=0 MB  
 bounds: [0.660530..0.660530] (0.000000) loss=0.540530, iterations=8793  
 here2  
 gosdt reported successful execution  
 training completed. 0.000/0.000/58.835 (user, system, wall), mem=0 MB  
 bounds: [0.740667..0.740667] (0.000000) loss=0.540667, iterations=8841  
 gosdt reported successful execution  
 training completed. 0.000/0.000/56.771 (user, system, wall), mem=0 MB  
 bounds: [0.749765..0.749765] (0.000000) loss=0.599765, iterations=8715  
 gosdt reported successful execution  
 training completed. 0.000/0.000/60.155 (user, system, wall), mem=0 MB  
 bounds: [0.730677..0.730677] (0.000000) loss=0.580677, iterations=8927  
 gosdt reported successful execution  
 training completed. 0.000/0.000/58.890 (user, system, wall), mem=0 MB  
 bounds: [0.737025..0.737025] (0.000000) loss=0.587025, iterations=8833  
 gosdt reported successful execution  
 training completed. 0.000/0.000/57.983 (user, system, wall), mem=0 MB  
 bounds: [0.737048..0.737048] (0.000000) loss=0.587048, iterations=8793  
 here2  
 gosdt reported successful execution

training completed. 0.000/0.000/65.000 (user, system, wall), mem=0 MB  
 bounds: [0.448669..0.448669] (0.000000) loss=0.408669, iterations=8841  
 gosdt reported successful execution  
 training completed. 0.000/0.000/63.043 (user, system, wall), mem=0 MB  
 bounds: [0.455893..0.455893] (0.000000) loss=0.415893, iterations=8715  
 gosdt reported successful execution  
 training completed. 0.000/0.000/66.315 (user, system, wall), mem=0 MB  
 bounds: [0.463711..0.463711] (0.000000) loss=0.423711, iterations=8927  
 gosdt reported successful execution  
 training completed. 0.000/0.000/64.389 (user, system, wall), mem=0 MB  
 bounds: [0.450309..0.450309] (0.000000) loss=0.410309, iterations=8833  
 gosdt reported successful execution  
 training completed. 0.000/0.000/64.187 (user, system, wall), mem=0 MB  
 bounds: [0.447217..0.447217] (0.000000) loss=0.407216, iterations=8793  
 here2  
 gosdt reported successful execution  
 training completed. 0.000/0.000/61.144 (user, system, wall), mem=0 MB  
 bounds: [0.489236..0.489236] (0.000000) loss=0.459236, iterations=8841  
 gosdt reported successful execution  
 training completed. 0.000/0.000/60.075 (user, system, wall), mem=0 MB  
 bounds: [0.498524..0.498524] (0.000000) loss=0.468524, iterations=8715  
 gosdt reported successful execution  
 training completed. 0.000/0.000/62.805 (user, system, wall), mem=0 MB  
 bounds: [0.505258..0.505258] (0.000000) loss=0.475258, iterations=8927  
 gosdt reported successful execution  
 training completed. 0.000/0.000/60.941 (user, system, wall), mem=0 MB  
 bounds: [0.498041..0.498041] (0.000000) loss=0.468041, iterations=8833  
 gosdt reported successful execution  
 training completed. 0.000/0.000/60.728 (user, system, wall), mem=0 MB  
 bounds: [0.498041..0.498041] (0.000000) loss=0.468041, iterations=8793  
 here2  
 gosdt reported successful execution  
 training completed. 0.000/0.000/57.974 (user, system, wall), mem=0 MB  
 bounds: [0.509236..0.509236] (0.000000) loss=0.459236, iterations=8841  
 gosdt reported successful execution  
 training completed. 0.000/0.000/56.349 (user, system, wall), mem=0 MB  
 bounds: [0.518524..0.518524] (0.000000) loss=0.468524, iterations=8715  
 gosdt reported successful execution  
 training completed. 0.000/0.000/59.704 (user, system, wall), mem=0 MB  
 bounds: [0.525258..0.525258] (0.000000) loss=0.475258, iterations=8927  
 gosdt reported successful execution  
 training completed. 0.000/0.000/57.878 (user, system, wall), mem=0 MB  
 bounds: [0.518041..0.518041] (0.000000) loss=0.468041, iterations=8833  
 gosdt reported successful execution  
 training completed. 0.000/0.000/57.090 (user, system, wall), mem=0 MB  
 bounds: [0.518041..0.518041] (0.000000) loss=0.468041, iterations=8793





here2

```

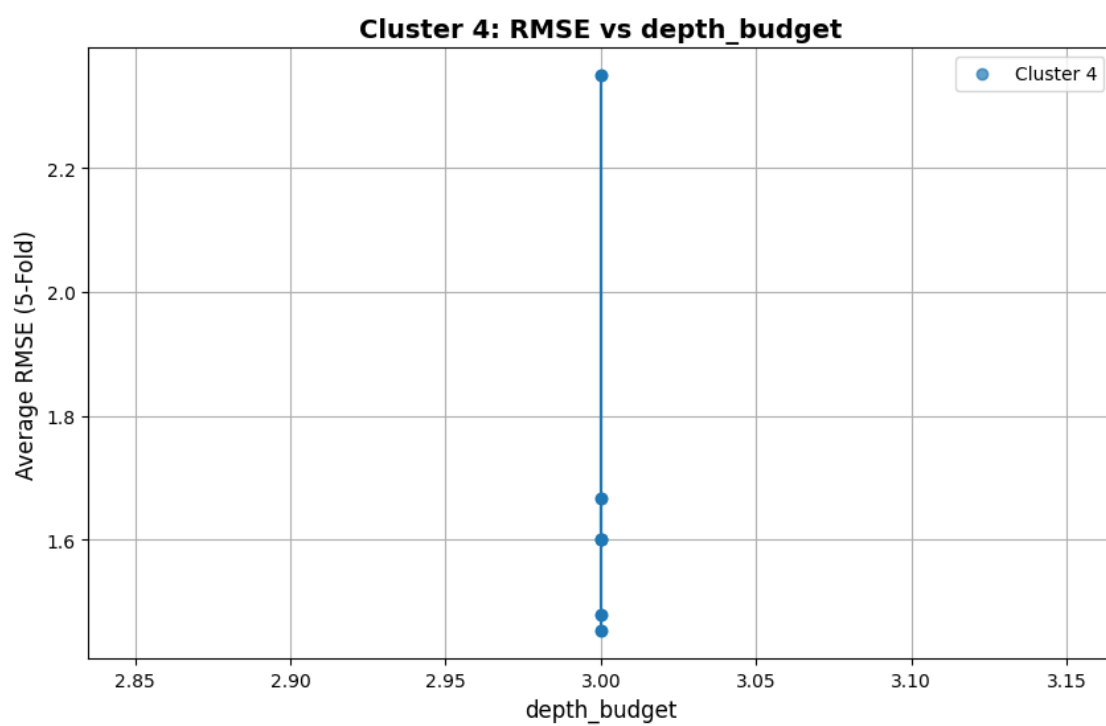
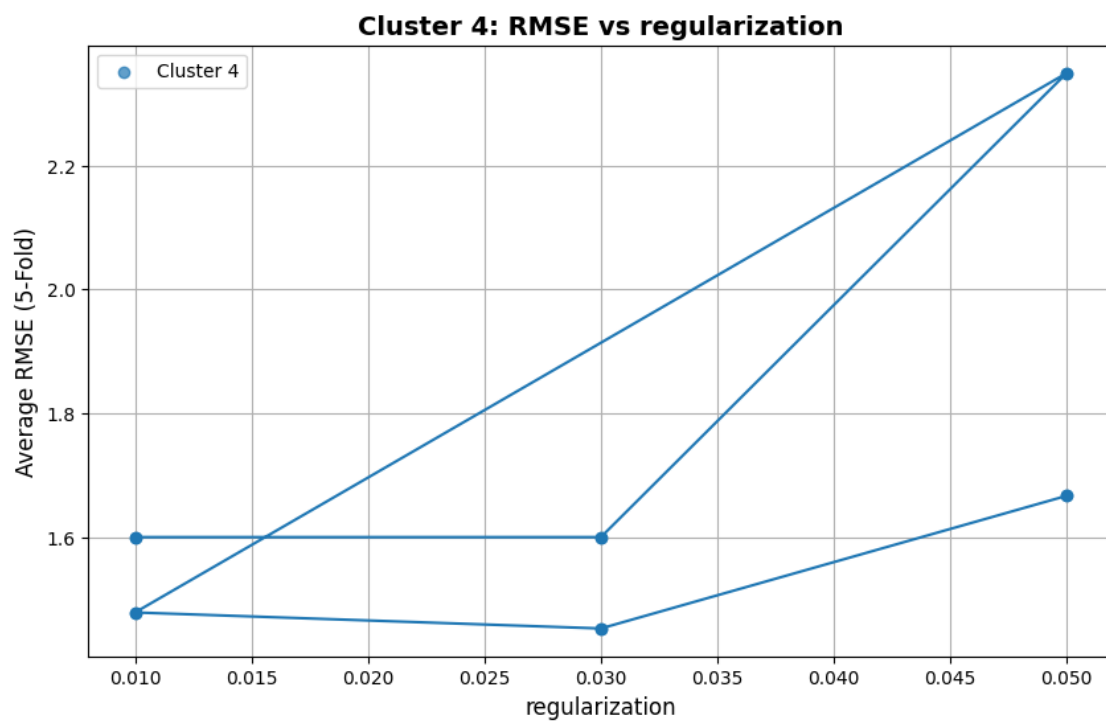
gosdt reported successful execution
training completed. 0.000/0.000/359.522 (user, system, wall), mem=0 MB
bounds: [0.662637..0.662637] (0.000000) loss=0.622637, iterations=16389
gosdt reported successful execution
training completed. 0.000/0.000/353.537 (user, system, wall), mem=0 MB
bounds: [0.664452..0.664452] (0.000000) loss=0.624452, iterations=16373
gosdt reported successful execution
training completed. 0.000/0.000/520.096 (user, system, wall), mem=0 MB
bounds: [0.659069..0.659069] (0.000000) loss=0.619069, iterations=16437
gosdt reported successful execution
training completed. 0.000/0.000/384.535 (user, system, wall), mem=0 MB
bounds: [0.655359..0.655359] (0.000000) loss=0.615359, iterations=16519
gosdt reported successful execution
training completed. 0.000/0.000/356.353 (user, system, wall), mem=0 MB
bounds: [0.662218..0.662218] (0.000000) loss=0.622218, iterations=16435
here2
gosdt reported successful execution
training completed. 0.000/0.000/344.811 (user, system, wall), mem=0 MB
bounds: [0.742637..0.742637] (0.000000) loss=0.622637, iterations=16389
gosdt reported successful execution
training completed. 0.000/0.000/339.927 (user, system, wall), mem=0 MB
bounds: [0.744452..0.744452] (0.000000) loss=0.624452, iterations=16373
gosdt reported successful execution
training completed. 0.000/0.000/342.613 (user, system, wall), mem=0 MB
bounds: [0.739069..0.739069] (0.000000) loss=0.619069, iterations=16437
gosdt reported successful execution
training completed. 0.000/0.000/346.635 (user, system, wall), mem=0 MB
bounds: [0.735359..0.735359] (0.000000) loss=0.615359, iterations=16519
gosdt reported successful execution
training completed. 0.000/0.000/342.636 (user, system, wall), mem=0 MB
bounds: [0.742218..0.742218] (0.000000) loss=0.622218, iterations=16435
here2
gosdt reported successful execution
training completed. 0.000/0.000/329.723 (user, system, wall), mem=0 MB
bounds: [0.818928..0.818928] (0.000000) loss=0.668928, iterations=16389
gosdt reported successful execution
training completed. 0.000/0.000/329.036 (user, system, wall), mem=0 MB
bounds: [0.818022..0.818022] (0.000000) loss=0.668022, iterations=16373
gosdt reported successful execution
training completed. 0.000/0.000/328.626 (user, system, wall), mem=0 MB
bounds: [0.816823..0.816823] (0.000000) loss=0.666822, iterations=16437
gosdt reported successful execution
training completed. 0.000/0.000/320.276 (user, system, wall), mem=0 MB
bounds: [0.814231..0.814231] (0.000000) loss=0.664231, iterations=16519
gosdt reported successful execution
training completed. 0.000/0.000/318.452 (user, system, wall), mem=0 MB
bounds: [0.816912..0.816912] (0.000000) loss=0.666912, iterations=16435
here2

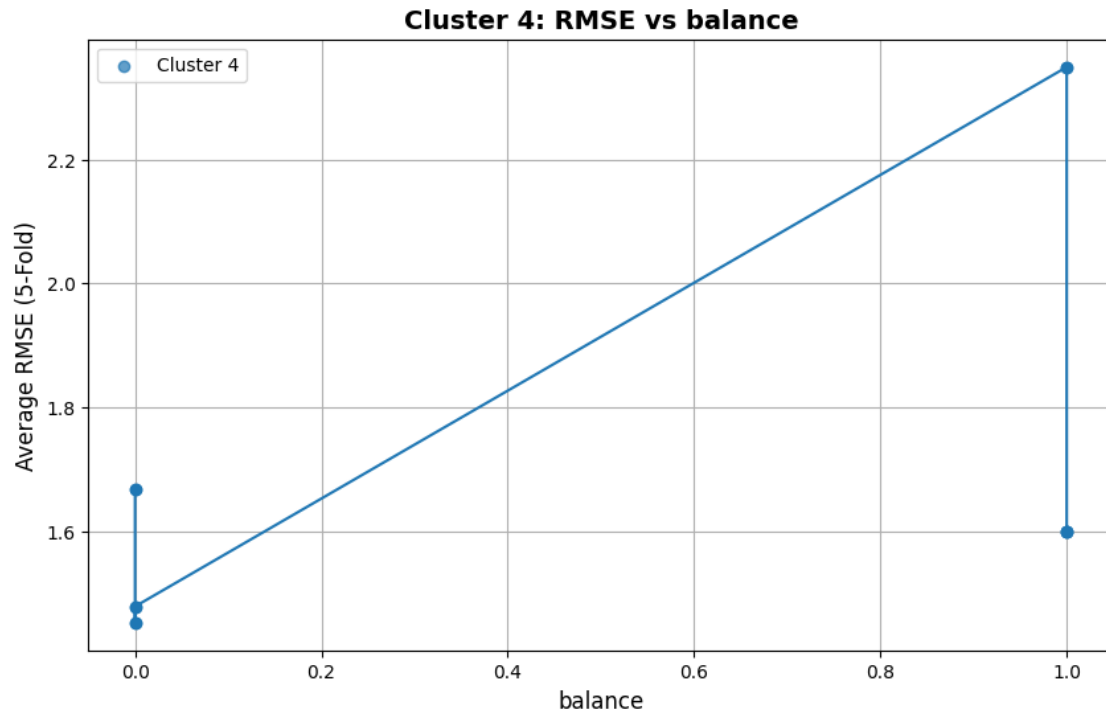
```

```

gosdt reported successful execution
training completed. 0.000/0.000/345.234 (user, system, wall), mem=0 MB
bounds: [0.614144..0.614144] (0.000000) loss=0.574144, iterations=16389
gosdt reported successful execution
training completed. 0.000/0.000/345.015 (user, system, wall), mem=0 MB
bounds: [0.614125..0.614125] (0.000000) loss=0.584125, iterations=16373
gosdt reported successful execution
training completed. 0.000/0.000/343.013 (user, system, wall), mem=0 MB
bounds: [0.612243..0.612243] (0.000000) loss=0.572243, iterations=16437
gosdt reported successful execution
training completed. 0.000/0.000/356.039 (user, system, wall), mem=0 MB
bounds: [0.608422..0.608422] (0.000000) loss=0.578422, iterations=16519
gosdt reported successful execution
training completed. 0.000/0.000/352.275 (user, system, wall), mem=0 MB
bounds: [0.617947..0.617947] (0.000000) loss=0.577947, iterations=16435
here2
gosdt reported successful execution
training completed. 0.000/0.000/336.850 (user, system, wall), mem=0 MB
bounds: [0.678403..0.678403] (0.000000) loss=0.588403, iterations=16389
gosdt reported successful execution
training completed. 0.000/0.000/331.074 (user, system, wall), mem=0 MB
bounds: [0.674068..0.674068] (0.000000) loss=0.614068, iterations=16373
gosdt reported successful execution
training completed. 0.000/0.000/334.572 (user, system, wall), mem=0 MB
bounds: [0.673650..0.673650] (0.000000) loss=0.583650, iterations=16437
gosdt reported successful execution
training completed. 0.000/0.000/330.132 (user, system, wall), mem=0 MB
bounds: [0.668422..0.668422] (0.000000) loss=0.578422, iterations=16519
gosdt reported successful execution
training completed. 0.000/0.000/326.898 (user, system, wall), mem=0 MB
bounds: [0.685532..0.685532] (0.000000) loss=0.595532, iterations=16435
here2
gosdt reported successful execution
training completed. 0.000/0.000/311.151 (user, system, wall), mem=0 MB
bounds: [0.720247..0.720247] (0.000000) loss=0.620247, iterations=16389
gosdt reported successful execution
training completed. 0.000/0.000/310.173 (user, system, wall), mem=0 MB
bounds: [0.714068..0.714068] (0.000000) loss=0.614068, iterations=16373
gosdt reported successful execution
training completed. 0.000/0.000/325.429 (user, system, wall), mem=0 MB
bounds: [0.716445..0.716445] (0.000000) loss=0.616445, iterations=16437
gosdt reported successful execution
training completed. 0.000/0.000/330.893 (user, system, wall), mem=0 MB
bounds: [0.711217..0.711217] (0.000000) loss=0.611217, iterations=16519
gosdt reported successful execution
training completed. 0.000/0.000/329.153 (user, system, wall), mem=0 MB
bounds: [0.729278..0.729278] (0.000000) loss=0.629278, iterations=16435

```





here2

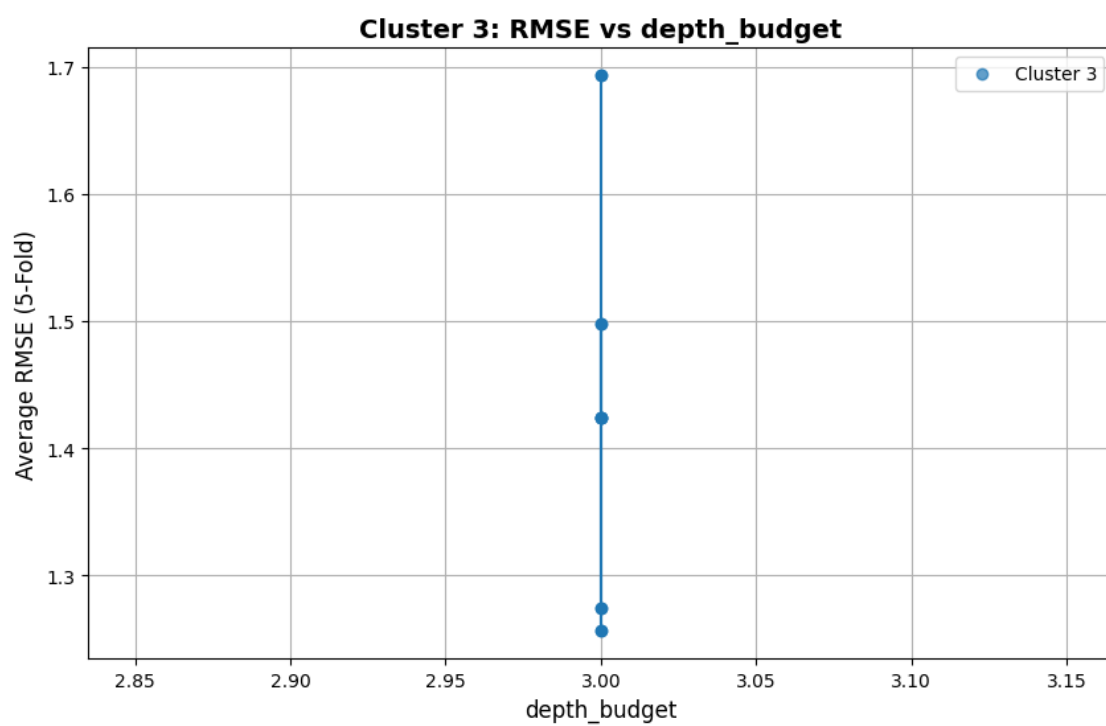
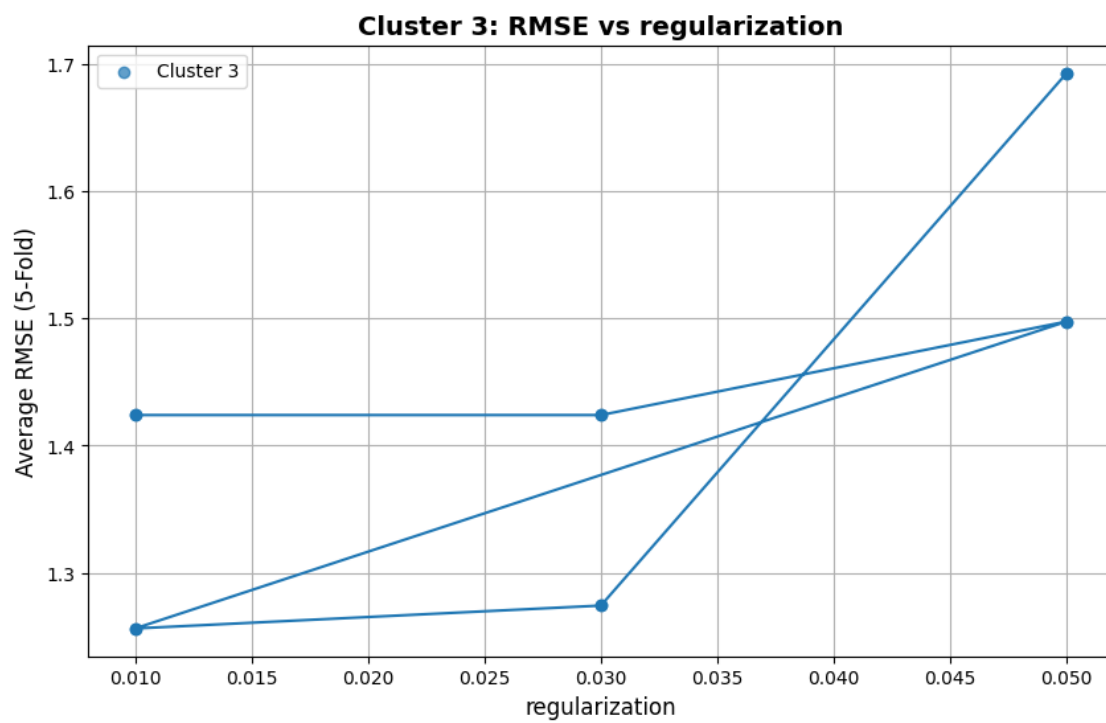


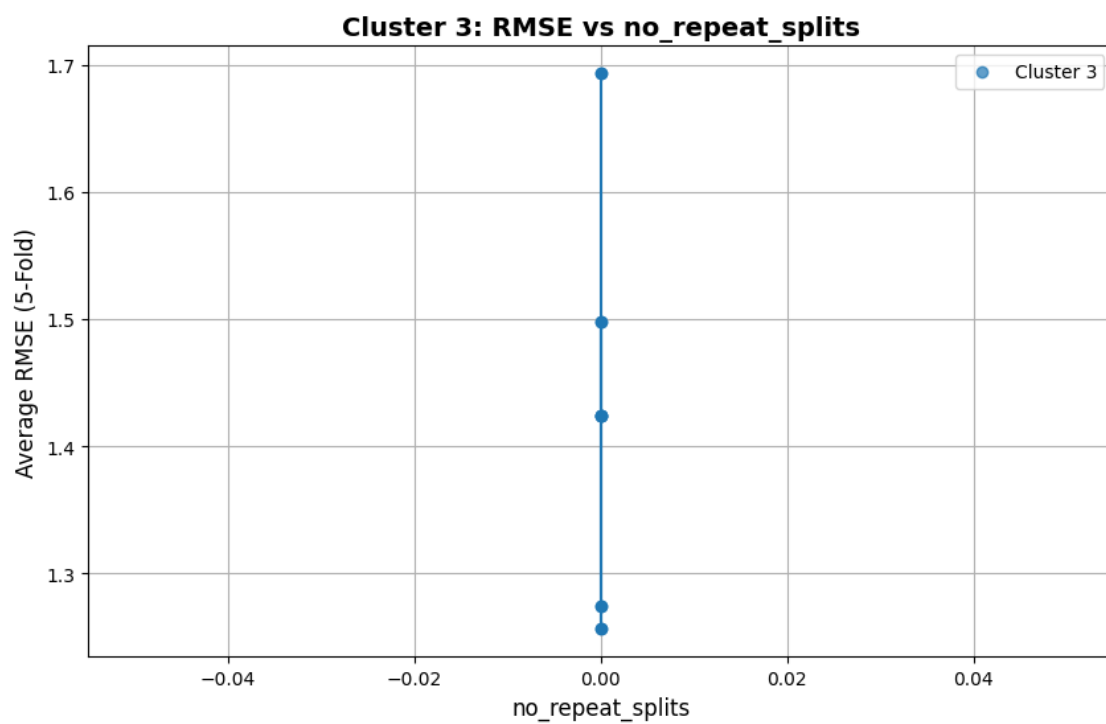
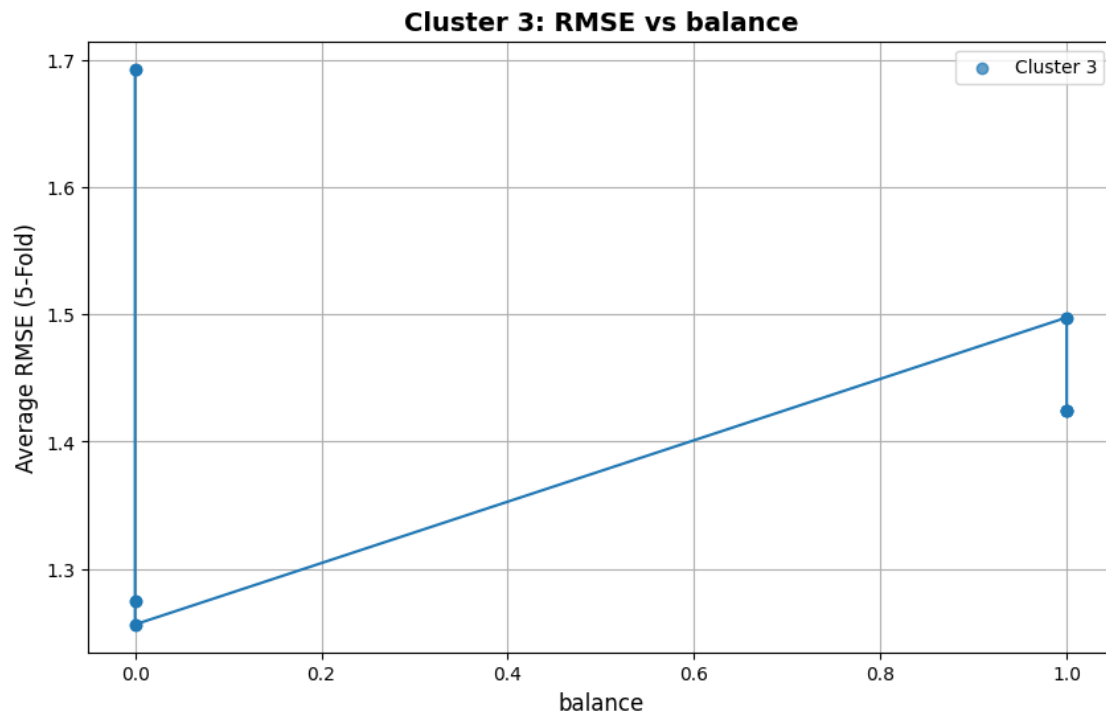
gosdt reported successful execution  
 training completed. 0.000/0.000/175.005 (user, system, wall), mem=0 MB  
 bounds: [0.645813..0.645813] (0.000000) loss=0.605813, iterations=13849  
 gosdt reported successful execution  
 training completed. 0.000/0.000/177.249 (user, system, wall), mem=0 MB  
 bounds: [0.638591..0.638591] (0.000000) loss=0.598591, iterations=13843  
 gosdt reported successful execution  
 training completed. 0.000/0.000/177.816 (user, system, wall), mem=0 MB  
 bounds: [0.643106..0.643106] (0.000000) loss=0.603106, iterations=13959  
 gosdt reported successful execution  
 training completed. 0.000/0.000/180.235 (user, system, wall), mem=0 MB  
 bounds: [0.650130..0.650130] (0.000000) loss=0.610130, iterations=13939  
 gosdt reported successful execution  
 training completed. 0.000/0.000/176.117 (user, system, wall), mem=0 MB  
 bounds: [0.641176..0.641176] (0.000000) loss=0.601176, iterations=13865  
 here2  
 gosdt reported successful execution  
 training completed. 0.000/0.000/169.333 (user, system, wall), mem=0 MB  
 bounds: [0.725813..0.725813] (0.000000) loss=0.605813, iterations=13849  
 gosdt reported successful execution  
 training completed. 0.000/0.000/169.896 (user, system, wall), mem=0 MB  
 bounds: [0.718591..0.718591] (0.000000) loss=0.598591, iterations=13843  
 gosdt reported successful execution  
 training completed. 0.000/0.000/170.029 (user, system, wall), mem=0 MB  
 bounds: [0.723106..0.723106] (0.000000) loss=0.603106, iterations=13959  
 gosdt reported successful execution  
 training completed. 0.000/0.000/171.859 (user, system, wall), mem=0 MB  
 bounds: [0.730130..0.730130] (0.000000) loss=0.610130, iterations=13939  
 gosdt reported successful execution  
 training completed. 0.000/0.000/168.090 (user, system, wall), mem=0 MB  
 bounds: [0.721176..0.721176] (0.000000) loss=0.601176, iterations=13865  
 here2  
 gosdt reported successful execution  
 training completed. 0.000/0.000/163.296 (user, system, wall), mem=0 MB  
 bounds: [0.790711..0.790711] (0.000000) loss=0.640711, iterations=13849  
 gosdt reported successful execution  
 training completed. 0.000/0.000/162.605 (user, system, wall), mem=0 MB  
 bounds: [0.779370..0.779370] (0.000000) loss=0.629370, iterations=13843  
 gosdt reported successful execution  
 training completed. 0.000/0.000/163.365 (user, system, wall), mem=0 MB  
 bounds: [0.791189..0.791189] (0.000000) loss=0.641189, iterations=13959  
 gosdt reported successful execution  
 training completed. 0.000/0.000/165.181 (user, system, wall), mem=0 MB  
 bounds: [0.795797..0.795797] (0.000000) loss=0.645797, iterations=13939  
 gosdt reported successful execution  
 training completed. 0.000/0.000/160.601 (user, system, wall), mem=0 MB  
 bounds: [0.787247..0.787247] (0.000000) loss=0.637247, iterations=13865  
 here2

```

gosdt reported successful execution
training completed. 0.000/0.000/176.816 (user, system, wall), mem=0 MB
bounds: [0.586089..0.586089] (0.000000) loss=0.546089, iterations=13849
gosdt reported successful execution
training completed. 0.000/0.000/178.896 (user, system, wall), mem=0 MB
bounds: [0.572449..0.572449] (0.000000) loss=0.532449, iterations=13843
gosdt reported successful execution
training completed. 0.000/0.000/192.339 (user, system, wall), mem=0 MB
bounds: [0.581521..0.581521] (0.000000) loss=0.541521, iterations=13959
gosdt reported successful execution
training completed. 0.000/0.000/185.695 (user, system, wall), mem=0 MB
bounds: [0.578032..0.578032] (0.000000) loss=0.538032, iterations=13939
gosdt reported successful execution
training completed. 0.000/0.000/179.071 (user, system, wall), mem=0 MB
bounds: [0.574543..0.574543] (0.000000) loss=0.534543, iterations=13865
here2
gosdt reported successful execution
training completed. 0.000/0.000/171.961 (user, system, wall), mem=0 MB
bounds: [0.652849..0.652849] (0.000000) loss=0.562849, iterations=13849
gosdt reported successful execution
training completed. 0.000/0.000/173.154 (user, system, wall), mem=0 MB
bounds: [0.641989..0.641989] (0.000000) loss=0.551989, iterations=13843
gosdt reported successful execution
training completed. 0.000/0.000/179.583 (user, system, wall), mem=0 MB
bounds: [0.651061..0.651061] (0.000000) loss=0.561061, iterations=13959
gosdt reported successful execution
training completed. 0.000/0.000/179.812 (user, system, wall), mem=0 MB
bounds: [0.644082..0.644082] (0.000000) loss=0.554082, iterations=13939
gosdt reported successful execution
training completed. 0.000/0.000/173.409 (user, system, wall), mem=0 MB
bounds: [0.639895..0.639895] (0.000000) loss=0.549895, iterations=13865
here2
gosdt reported successful execution
training completed. 0.000/0.000/168.239 (user, system, wall), mem=0 MB
bounds: [0.706145..0.706145] (0.000000) loss=0.606145, iterations=13849
gosdt reported successful execution
training completed. 0.000/0.000/167.676 (user, system, wall), mem=0 MB
bounds: [0.694557..0.694557] (0.000000) loss=0.594557, iterations=13843
gosdt reported successful execution
training completed. 0.000/0.000/166.741 (user, system, wall), mem=0 MB
bounds: [0.702233..0.702233] (0.000000) loss=0.602233, iterations=13959
gosdt reported successful execution
training completed. 0.000/0.000/168.085 (user, system, wall), mem=0 MB
bounds: [0.695255..0.695255] (0.000000) loss=0.595255, iterations=13939
gosdt reported successful execution
training completed. 0.000/0.000/166.758 (user, system, wall), mem=0 MB
bounds: [0.688974..0.688974] (0.000000) loss=0.588974, iterations=13865

```





here2

```

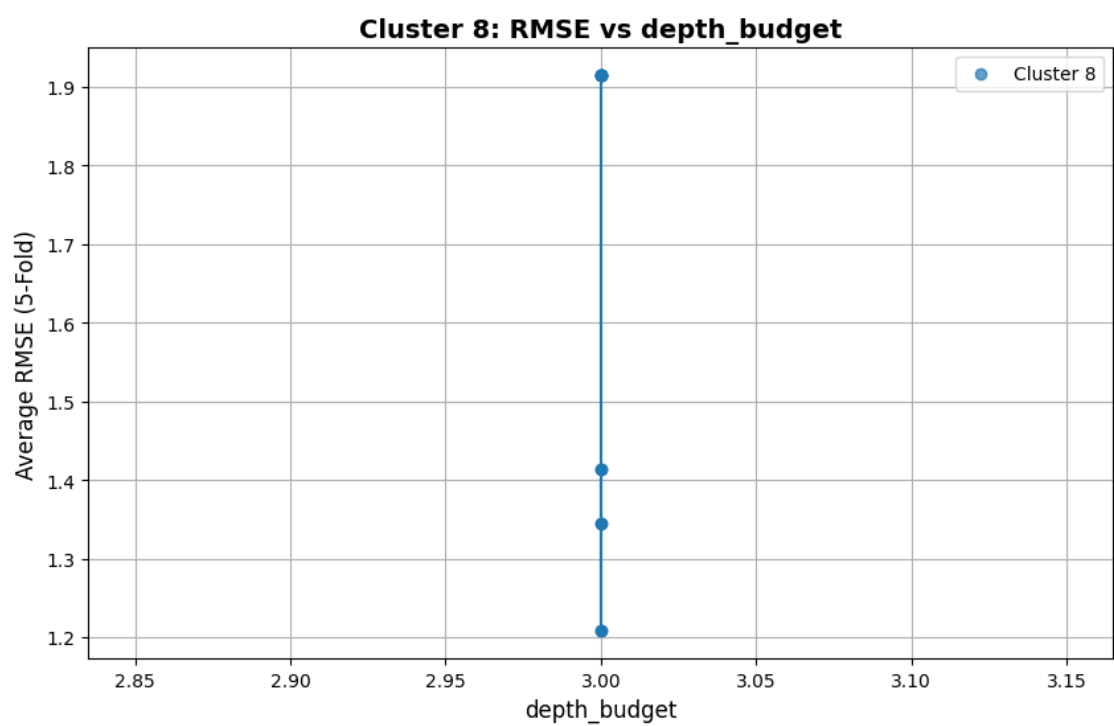
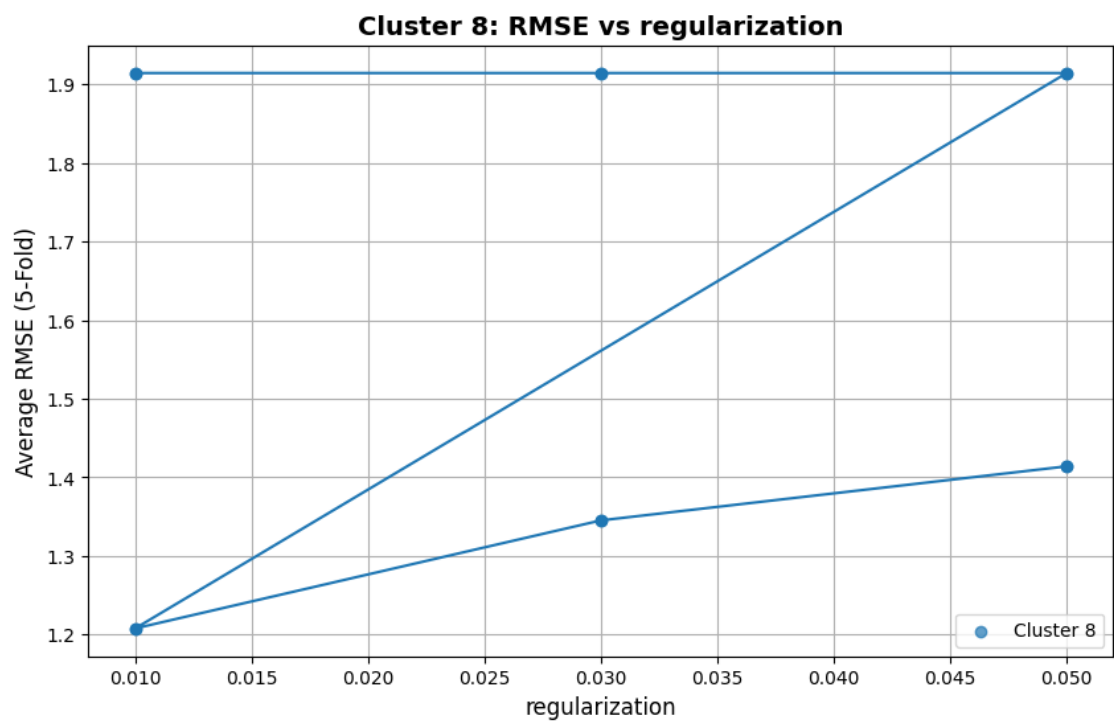
gosdt reported successful execution
training completed. 0.000/0.000/20.383 (user, system, wall), mem=0 MB
bounds: [0.610256..0.610256] (0.000000) loss=0.570256, iterations=5437
gosdt reported successful execution
training completed. 0.000/0.000/20.077 (user, system, wall), mem=0 MB
bounds: [0.629442..0.629442] (0.000000) loss=0.589442, iterations=5425
gosdt reported successful execution
training completed. 0.000/0.000/19.649 (user, system, wall), mem=0 MB
bounds: [0.628990..0.628990] (0.000000) loss=0.588990, iterations=5373
gosdt reported successful execution
training completed. 0.000/0.000/19.001 (user, system, wall), mem=0 MB
bounds: [0.626920..0.626920] (0.000000) loss=0.586920, iterations=5333
gosdt reported successful execution
training completed. 0.000/0.000/20.056 (user, system, wall), mem=0 MB
bounds: [0.620148..0.620148] (0.000000) loss=0.580148, iterations=5501
here2
gosdt reported successful execution
training completed. 0.000/0.000/18.943 (user, system, wall), mem=0 MB
bounds: [0.690256..0.690256] (0.000000) loss=0.570256, iterations=5437
gosdt reported successful execution
training completed. 0.000/0.000/19.175 (user, system, wall), mem=0 MB
bounds: [0.709442..0.709442] (0.000000) loss=0.589442, iterations=5425
gosdt reported successful execution
training completed. 0.000/0.000/18.776 (user, system, wall), mem=0 MB
bounds: [0.708990..0.708990] (0.000000) loss=0.588990, iterations=5373
gosdt reported successful execution
training completed. 0.000/0.000/18.553 (user, system, wall), mem=0 MB
bounds: [0.706920..0.706920] (0.000000) loss=0.586920, iterations=5333
gosdt reported successful execution
training completed. 0.000/0.000/19.874 (user, system, wall), mem=0 MB
bounds: [0.700148..0.700148] (0.000000) loss=0.580148, iterations=5501
here2
gosdt reported successful execution
training completed. 0.000/0.000/18.494 (user, system, wall), mem=0 MB
bounds: [0.770256..0.770256] (0.000000) loss=0.570256, iterations=5437
gosdt reported successful execution
training completed. 0.000/0.000/18.427 (user, system, wall), mem=0 MB
bounds: [0.789442..0.789442] (0.000000) loss=0.589442, iterations=5425
gosdt reported successful execution
training completed. 0.000/0.000/18.182 (user, system, wall), mem=0 MB
bounds: [0.788990..0.788990] (0.000000) loss=0.588990, iterations=5373
gosdt reported successful execution
training completed. 0.000/0.000/17.811 (user, system, wall), mem=0 MB
bounds: [0.786920..0.786920] (0.000000) loss=0.586920, iterations=5333
gosdt reported successful execution
training completed. 0.000/0.000/18.870 (user, system, wall), mem=0 MB
bounds: [0.780148..0.780148] (0.000000) loss=0.580148, iterations=5501
here2

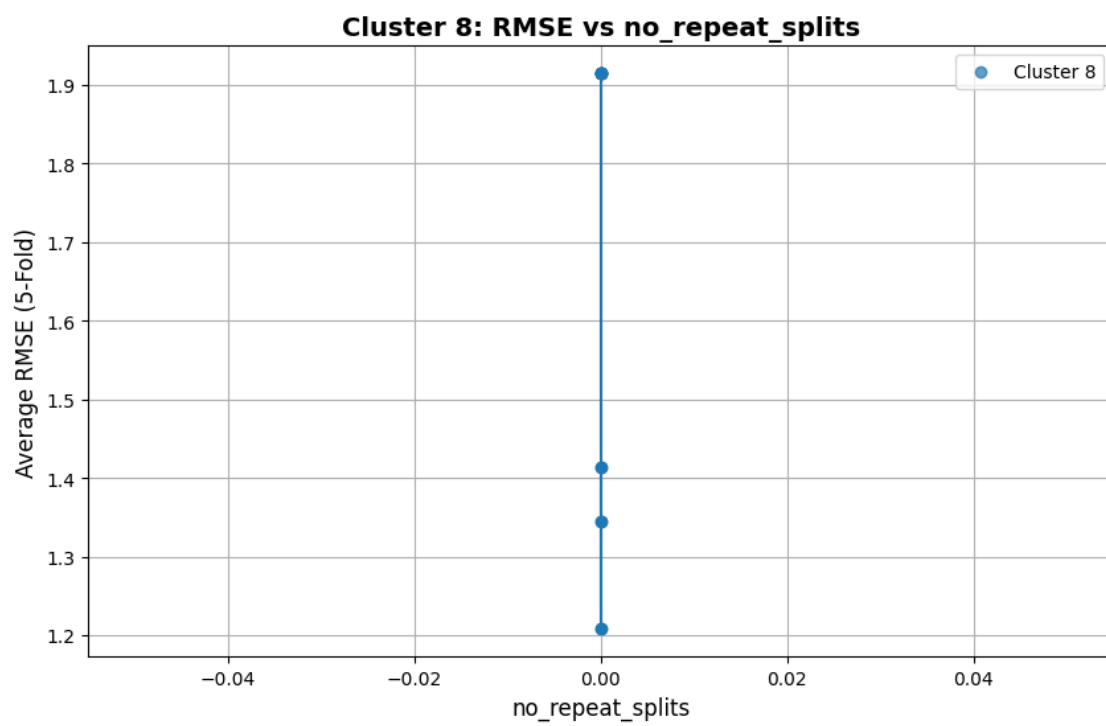
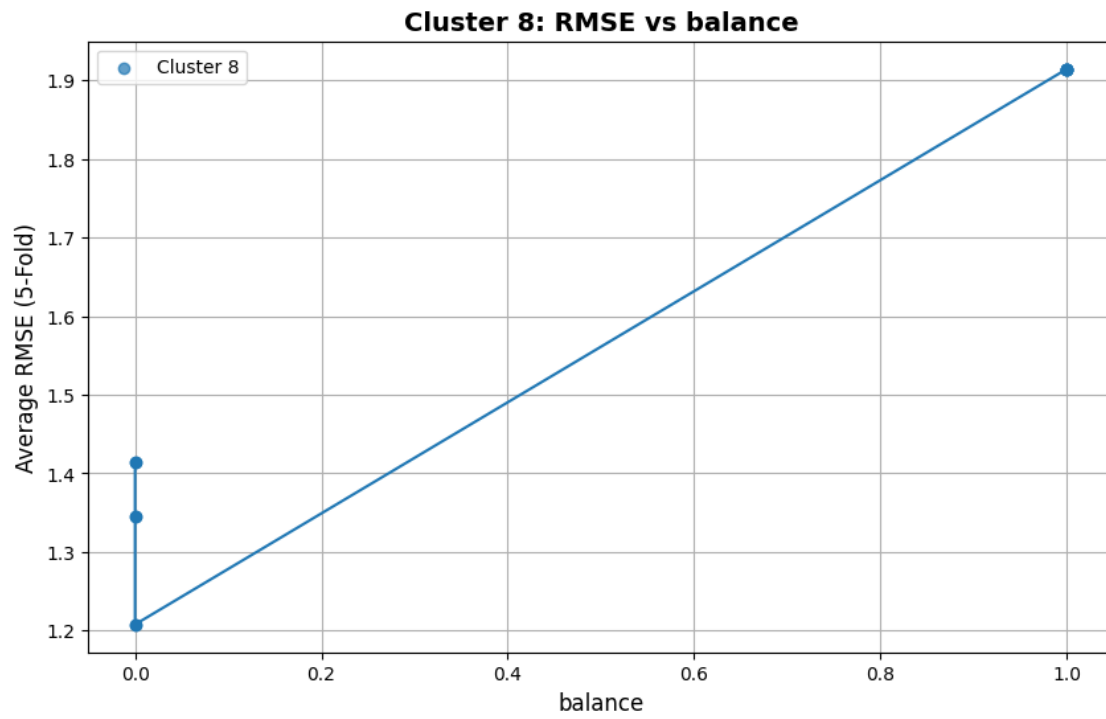
```

```

gosdt reported successful execution
training completed. 0.000/0.000/19.889 (user, system, wall), mem=0 MB
bounds: [0.578298..0.578298] (0.000000) loss=0.538298, iterations=5437
gosdt reported successful execution
training completed. 0.000/0.000/19.683 (user, system, wall), mem=0 MB
bounds: [0.593191..0.593191] (0.000000) loss=0.553191, iterations=5425
gosdt reported successful execution
training completed. 0.000/0.000/19.271 (user, system, wall), mem=0 MB
bounds: [0.563404..0.563404] (0.000000) loss=0.523404, iterations=5373
gosdt reported successful execution
training completed. 0.000/0.000/18.992 (user, system, wall), mem=0 MB
bounds: [0.568662..0.568662] (0.000000) loss=0.528662, iterations=5333
gosdt reported successful execution
training completed. 0.000/0.000/20.187 (user, system, wall), mem=0 MB
bounds: [0.581401..0.581401] (0.000000) loss=0.541401, iterations=5501
here2
gosdt reported successful execution
training completed. 0.000/0.000/19.193 (user, system, wall), mem=0 MB
bounds: [0.653830..0.653830] (0.000000) loss=0.563830, iterations=5437
gosdt reported successful execution
training completed. 0.000/0.000/18.848 (user, system, wall), mem=0 MB
bounds: [0.670851..0.670851] (0.000000) loss=0.580851, iterations=5425
gosdt reported successful execution
training completed. 0.000/0.000/18.660 (user, system, wall), mem=0 MB
bounds: [0.636808..0.636808] (0.000000) loss=0.546808, iterations=5373
gosdt reported successful execution
training completed. 0.000/0.000/18.361 (user, system, wall), mem=0 MB
bounds: [0.646263..0.646263] (0.000000) loss=0.556263, iterations=5333
gosdt reported successful execution
training completed. 0.000/0.000/19.491 (user, system, wall), mem=0 MB
bounds: [0.661401..0.661401] (0.000000) loss=0.541401, iterations=5501
here2
gosdt reported successful execution
training completed. 0.000/0.000/18.311 (user, system, wall), mem=0 MB
bounds: [0.713830..0.713830] (0.000000) loss=0.563830, iterations=5437
gosdt reported successful execution
training completed. 0.000/0.000/18.416 (user, system, wall), mem=0 MB
bounds: [0.730851..0.730851] (0.000000) loss=0.580851, iterations=5425
gosdt reported successful execution
training completed. 0.000/0.000/17.829 (user, system, wall), mem=0 MB
bounds: [0.696808..0.696808] (0.000000) loss=0.546808, iterations=5373
gosdt reported successful execution
training completed. 0.000/0.000/17.443 (user, system, wall), mem=0 MB
bounds: [0.706263..0.706263] (0.000000) loss=0.556263, iterations=5333
gosdt reported successful execution
training completed. 0.000/0.000/18.538 (user, system, wall), mem=0 MB
bounds: [0.723248..0.723248] (0.000000) loss=0.573248, iterations=5501

```





here2



```

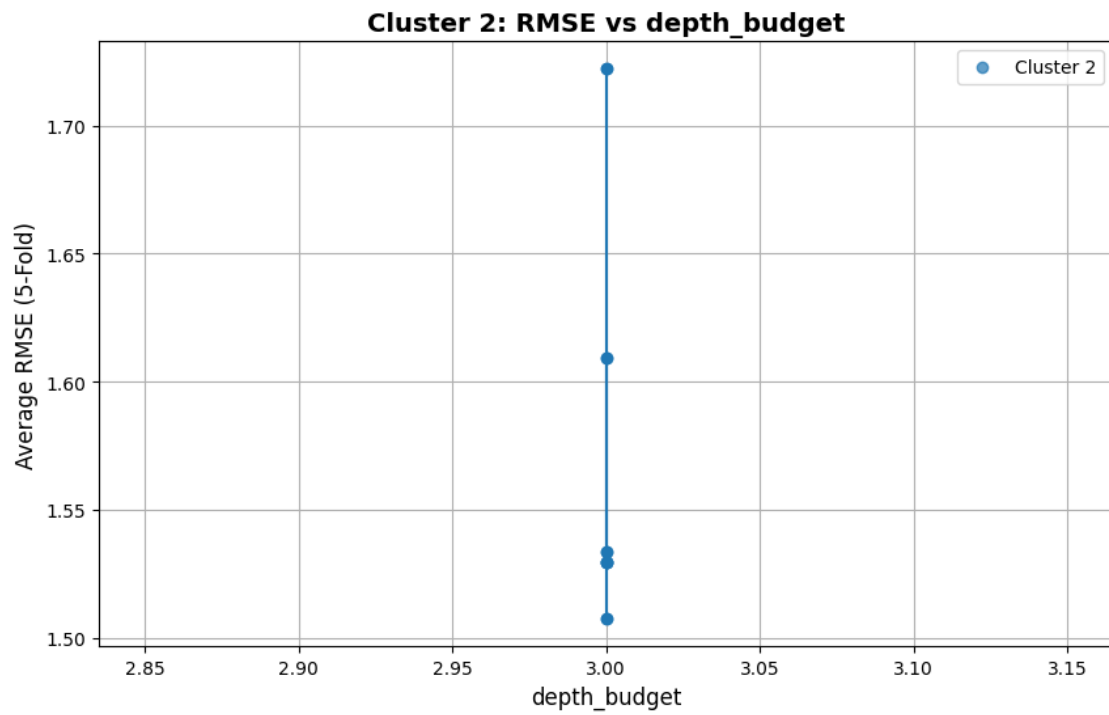
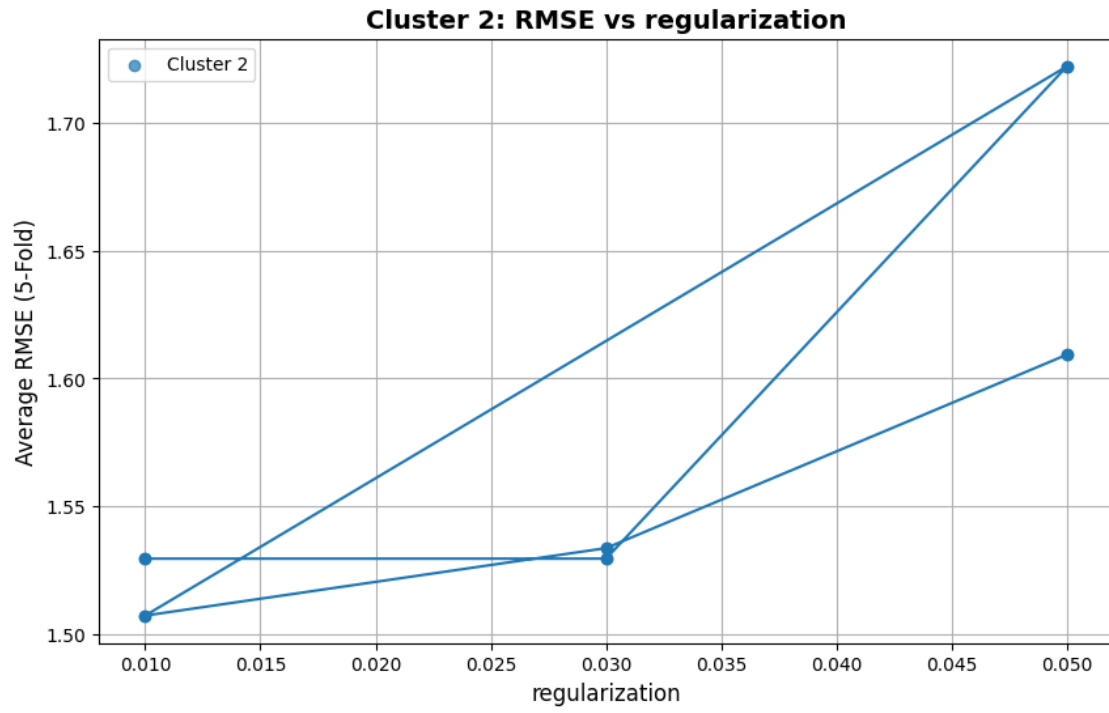
gosdt reported successful execution
training completed. 0.000/0.000/108.649 (user, system, wall), mem=0 MB
bounds: [0.635980..0.635980] (0.000000) loss=0.595980, iterations=11601
gosdt reported successful execution
training completed. 0.000/0.000/109.780 (user, system, wall), mem=0 MB
bounds: [0.626436..0.626436] (0.000000) loss=0.586436, iterations=11563
gosdt reported successful execution
training completed. 0.000/0.000/109.891 (user, system, wall), mem=0 MB
bounds: [0.641345..0.641345] (0.000000) loss=0.601345, iterations=11569
gosdt reported successful execution
training completed. 0.000/0.000/110.044 (user, system, wall), mem=0 MB
bounds: [0.636885..0.636885] (0.000000) loss=0.596885, iterations=11639
gosdt reported successful execution
training completed. 0.000/0.000/110.502 (user, system, wall), mem=0 MB
bounds: [0.621443..0.621443] (0.000000) loss=0.581443, iterations=11609
here2
gosdt reported successful execution
training completed. 0.000/0.000/104.340 (user, system, wall), mem=0 MB
bounds: [0.715980..0.715980] (0.000000) loss=0.595980, iterations=11601
gosdt reported successful execution
training completed. 0.000/0.000/104.433 (user, system, wall), mem=0 MB
bounds: [0.706436..0.706436] (0.000000) loss=0.586436, iterations=11563
gosdt reported successful execution
training completed. 0.000/0.000/103.742 (user, system, wall), mem=0 MB
bounds: [0.721345..0.721345] (0.000000) loss=0.601345, iterations=11569
gosdt reported successful execution
training completed. 0.000/0.000/105.272 (user, system, wall), mem=0 MB
bounds: [0.716885..0.716885] (0.000000) loss=0.596885, iterations=11639
gosdt reported successful execution
training completed. 0.000/0.000/104.336 (user, system, wall), mem=0 MB
bounds: [0.701443..0.701443] (0.000000) loss=0.581443, iterations=11609
here2
gosdt reported successful execution
training completed. 0.000/0.000/101.072 (user, system, wall), mem=0 MB
bounds: [0.792037..0.792037] (0.000000) loss=0.642037, iterations=11601
gosdt reported successful execution
training completed. 0.000/0.000/99.740 (user, system, wall), mem=0 MB
bounds: [0.786436..0.786436] (0.000000) loss=0.586436, iterations=11563
gosdt reported successful execution
training completed. 0.000/0.000/100.218 (user, system, wall), mem=0 MB
bounds: [0.791179..0.791179] (0.000000) loss=0.641179, iterations=11569
gosdt reported successful execution
training completed. 0.000/0.000/100.810 (user, system, wall), mem=0 MB
bounds: [0.789544..0.789544] (0.000000) loss=0.639544, iterations=11639
gosdt reported successful execution
training completed. 0.000/0.000/101.061 (user, system, wall), mem=0 MB
bounds: [0.781443..0.781443] (0.000000) loss=0.581443, iterations=11609
here2

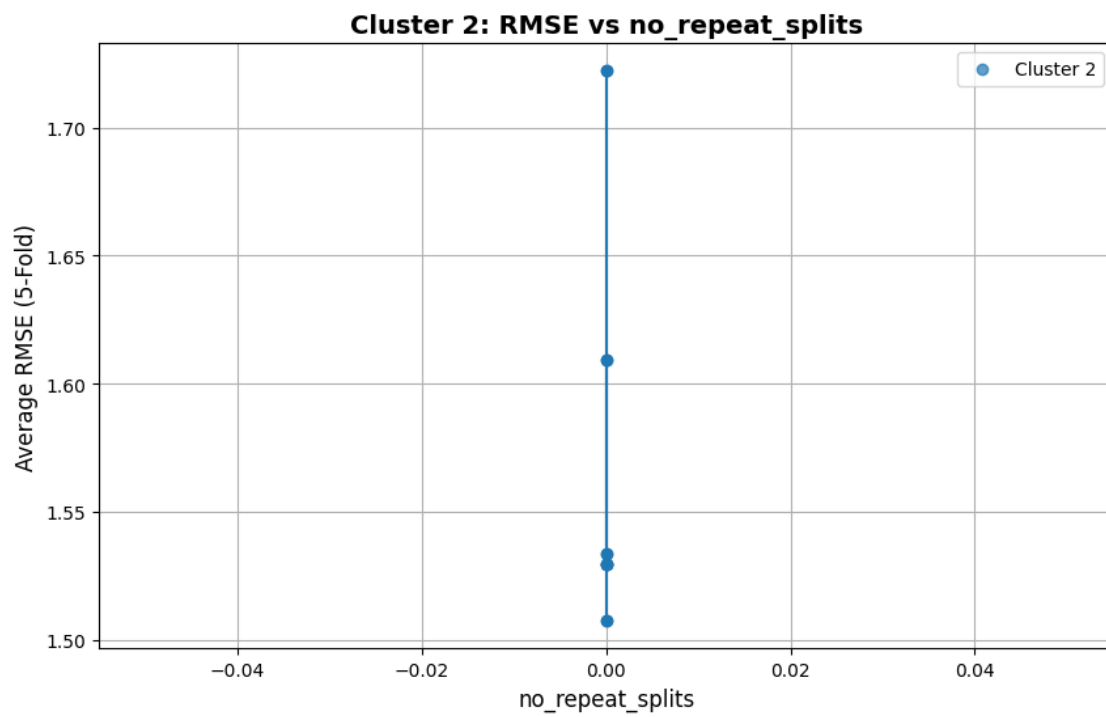
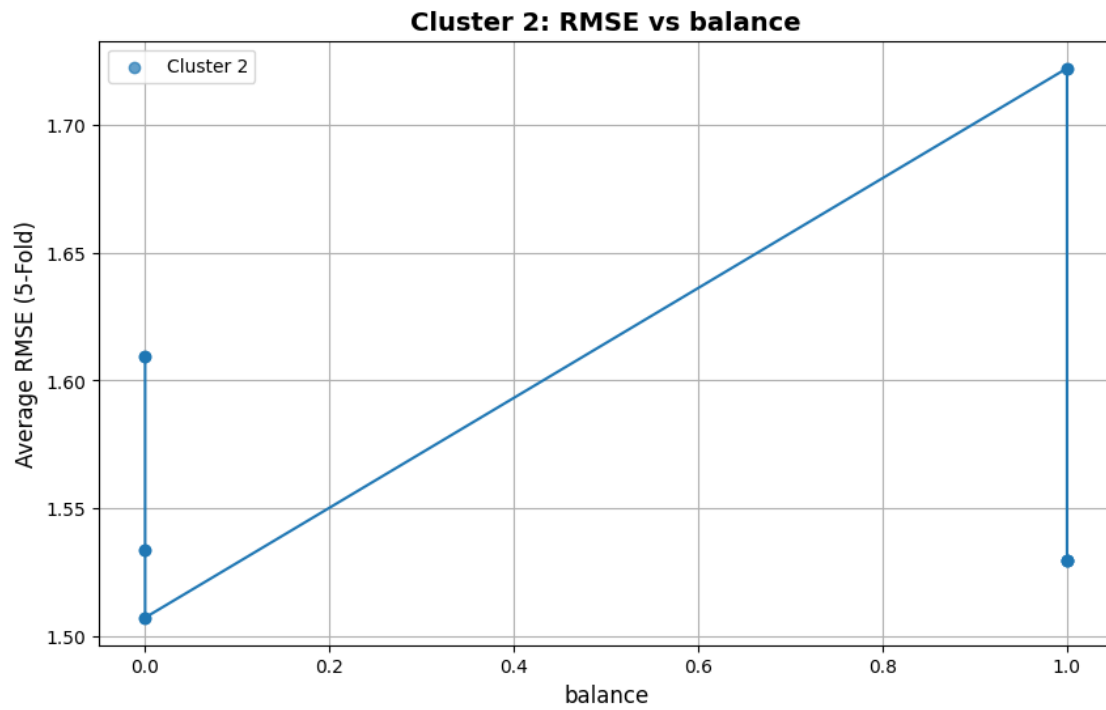
```

```

gosdt reported successful execution
training completed. 0.000/0.000/109.910 (user, system, wall), mem=0 MB
bounds: [0.670996..0.670996] (0.000000) loss=0.630996, iterations=11601
gosdt reported successful execution
training completed. 0.000/0.000/111.719 (user, system, wall), mem=0 MB
bounds: [0.660849..0.660849] (0.000000) loss=0.620849, iterations=11563
gosdt reported successful execution
training completed. 0.000/0.000/110.299 (user, system, wall), mem=0 MB
bounds: [0.669151..0.669151] (0.000000) loss=0.629151, iterations=11569
gosdt reported successful execution
training completed. 0.000/0.000/112.865 (user, system, wall), mem=0 MB
bounds: [0.676531..0.676531] (0.000000) loss=0.636531, iterations=11639
gosdt reported successful execution
training completed. 0.000/0.000/110.903 (user, system, wall), mem=0 MB
bounds: [0.662694..0.662694] (0.000000) loss=0.622694, iterations=11609
here2
gosdt reported successful execution
training completed. 0.000/0.000/105.380 (user, system, wall), mem=0 MB
bounds: [0.750996..0.750996] (0.000000) loss=0.630996, iterations=11601
gosdt reported successful execution
training completed. 0.000/0.000/104.142 (user, system, wall), mem=0 MB
bounds: [0.740849..0.740849] (0.000000) loss=0.620849, iterations=11563
gosdt reported successful execution
training completed. 0.000/0.000/105.599 (user, system, wall), mem=0 MB
bounds: [0.749151..0.749151] (0.000000) loss=0.629151, iterations=11569
gosdt reported successful execution
training completed. 0.000/0.000/108.077 (user, system, wall), mem=0 MB
bounds: [0.751439..0.751439] (0.000000) loss=0.661439, iterations=11639
gosdt reported successful execution
training completed. 0.000/0.000/106.992 (user, system, wall), mem=0 MB
bounds: [0.742694..0.742694] (0.000000) loss=0.622694, iterations=11609
here2
gosdt reported successful execution
training completed. 0.000/0.000/100.797 (user, system, wall), mem=0 MB
bounds: [0.807565..0.807565] (0.000000) loss=0.707565, iterations=11601
gosdt reported successful execution
training completed. 0.000/0.000/100.198 (user, system, wall), mem=0 MB
bounds: [0.795572..0.795572] (0.000000) loss=0.695572, iterations=11563
gosdt reported successful execution
training completed. 0.000/0.000/105.428 (user, system, wall), mem=0 MB
bounds: [0.790959..0.790959] (0.000000) loss=0.690959, iterations=11569
gosdt reported successful execution
training completed. 0.000/0.000/103.709 (user, system, wall), mem=0 MB
bounds: [0.796494..0.796494] (0.000000) loss=0.696494, iterations=11639
gosdt reported successful execution
training completed. 0.000/0.000/104.945 (user, system, wall), mem=0 MB
bounds: [0.803875..0.803875] (0.000000) loss=0.703875, iterations=11609

```





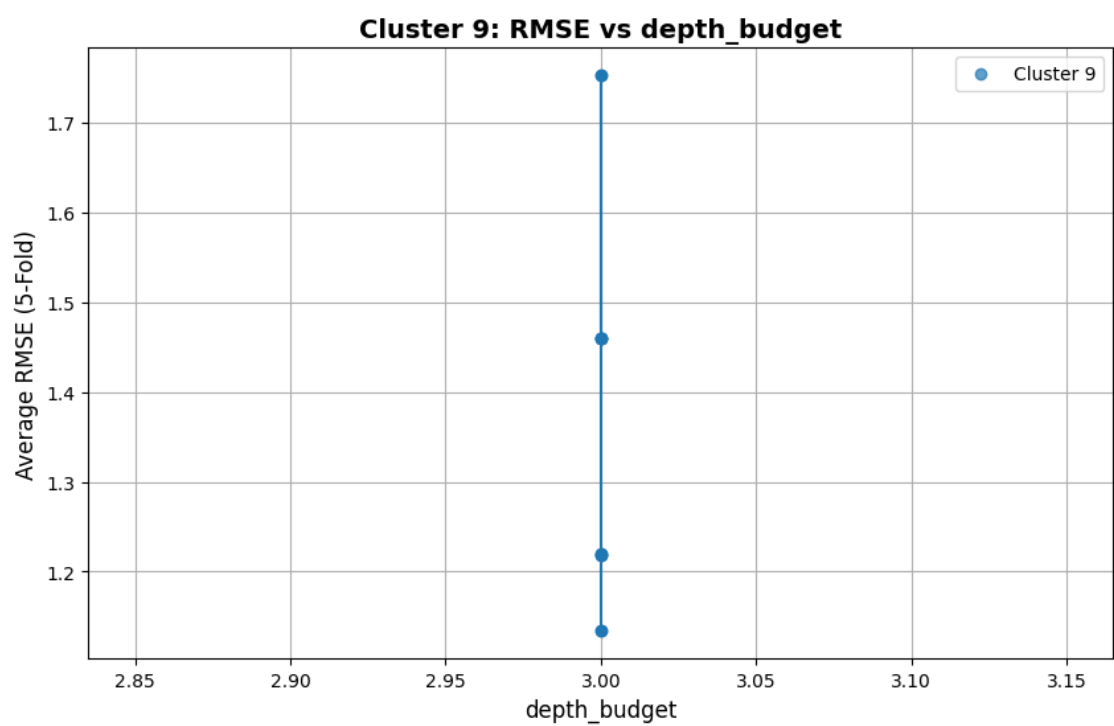
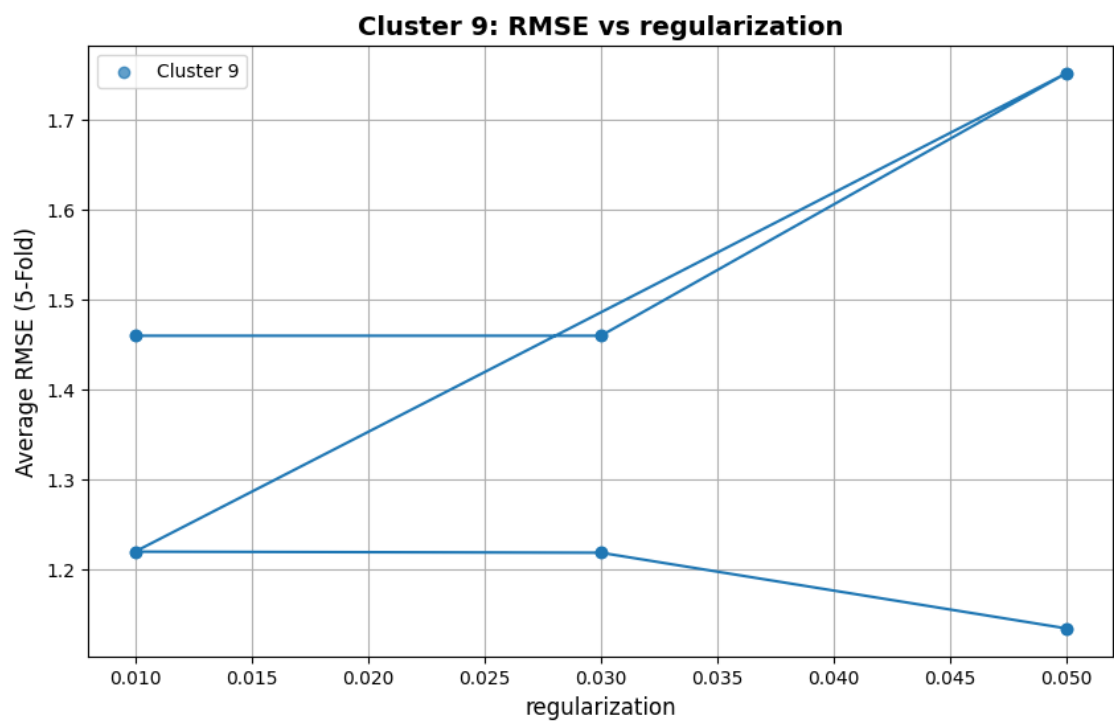
here2

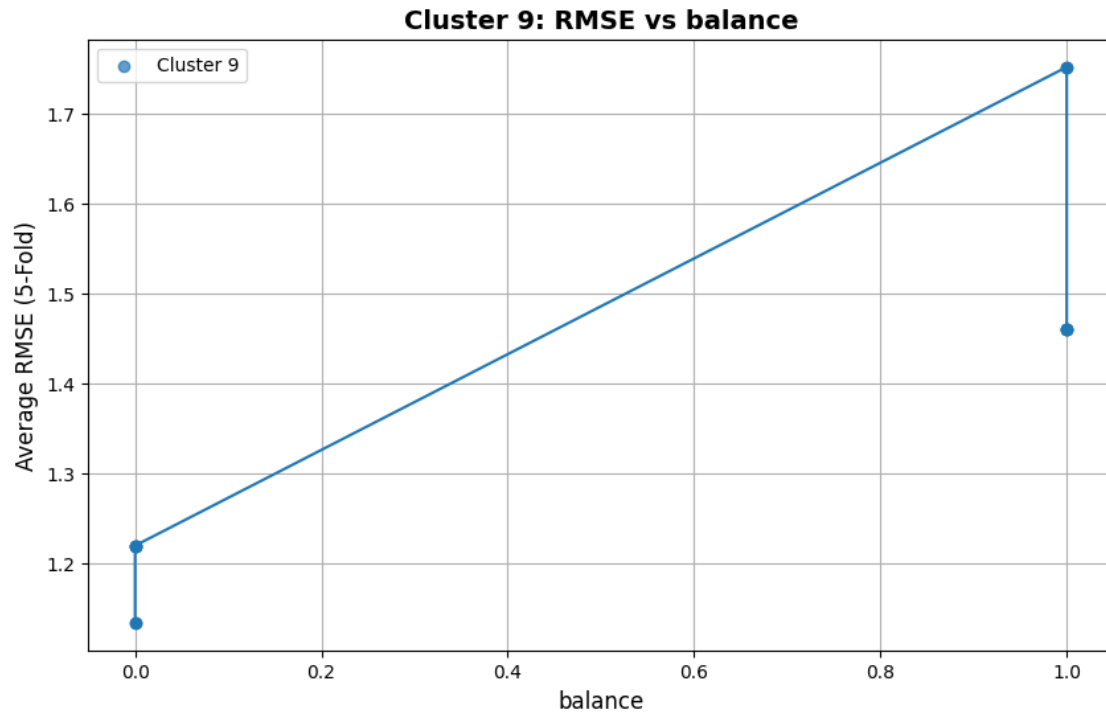
gosdt reported successful execution  
 training completed. 0.000/0.000/36.856 (user, system, wall), mem=0 MB  
 bounds: [0.607933..0.607933] (0.000000) loss=0.567933, iterations=7607  
 gosdt reported successful execution  
 training completed. 0.000/0.000/36.674 (user, system, wall), mem=0 MB  
 bounds: [0.618885..0.618885] (0.000000) loss=0.578885, iterations=7595  
 gosdt reported successful execution  
 training completed. 0.000/0.000/37.329 (user, system, wall), mem=0 MB  
 bounds: [0.620222..0.620222] (0.000000) loss=0.580222, iterations=7569  
 gosdt reported successful execution  
 training completed. 0.000/0.000/36.848 (user, system, wall), mem=0 MB  
 bounds: [0.623358..0.623358] (0.000000) loss=0.583358, iterations=7623  
 gosdt reported successful execution  
 training completed. 0.000/0.000/35.202 (user, system, wall), mem=0 MB  
 bounds: [0.615638..0.615638] (0.000000) loss=0.575638, iterations=7553  
 here2  
 gosdt reported successful execution  
 training completed. 0.000/0.000/34.810 (user, system, wall), mem=0 MB  
 bounds: [0.687933..0.687933] (0.000000) loss=0.567933, iterations=7607  
 gosdt reported successful execution  
 training completed. 0.000/0.000/34.709 (user, system, wall), mem=0 MB  
 bounds: [0.698885..0.698885] (0.000000) loss=0.578885, iterations=7595  
 gosdt reported successful execution  
 training completed. 0.000/0.000/35.076 (user, system, wall), mem=0 MB  
 bounds: [0.700222..0.700222] (0.000000) loss=0.580222, iterations=7569  
 gosdt reported successful execution  
 training completed. 0.000/0.000/34.014 (user, system, wall), mem=0 MB  
 bounds: [0.703358..0.703358] (0.000000) loss=0.583358, iterations=7623  
 gosdt reported successful execution  
 training completed. 0.000/0.000/33.468 (user, system, wall), mem=0 MB  
 bounds: [0.695638..0.695638] (0.000000) loss=0.575638, iterations=7553  
 here2  
 gosdt reported successful execution  
 training completed. 0.000/0.000/32.636 (user, system, wall), mem=0 MB  
 bounds: [0.759748..0.759748] (0.000000) loss=0.609748, iterations=7607  
 gosdt reported successful execution  
 training completed. 0.000/0.000/32.437 (user, system, wall), mem=0 MB  
 bounds: [0.774024..0.774024] (0.000000) loss=0.624024, iterations=7595  
 gosdt reported successful execution  
 training completed. 0.000/0.000/33.560 (user, system, wall), mem=0 MB  
 bounds: [0.770151..0.770151] (0.000000) loss=0.620151, iterations=7569  
 gosdt reported successful execution  
 training completed. 0.000/0.000/32.480 (user, system, wall), mem=0 MB  
 bounds: [0.771296..0.771296] (0.000000) loss=0.671296, iterations=7623  
 gosdt reported successful execution  
 training completed. 0.000/0.000/32.025 (user, system, wall), mem=0 MB  
 bounds: [0.774456..0.774456] (0.000000) loss=0.624456, iterations=7553  
 here2

```

gosdt reported successful execution
training completed. 0.000/0.000/35.179 (user, system, wall), mem=0 MB
bounds: [0.632150..0.632150] (0.000000) loss=0.592150, iterations=7607
gosdt reported successful execution
training completed. 0.000/0.000/36.260 (user, system, wall), mem=0 MB
bounds: [0.616792..0.616792] (0.000000) loss=0.576792, iterations=7595
gosdt reported successful execution
training completed. 0.000/0.000/37.627 (user, system, wall), mem=0 MB
bounds: [0.615085..0.615085] (0.000000) loss=0.575085, iterations=7569
gosdt reported successful execution
training completed. 0.000/0.000/35.035 (user, system, wall), mem=0 MB
bounds: [0.620920..0.620920] (0.000000) loss=0.580920, iterations=7623
gosdt reported successful execution
training completed. 0.000/0.000/34.465 (user, system, wall), mem=0 MB
bounds: [0.614106..0.614106] (0.000000) loss=0.574106, iterations=7553
here2
gosdt reported successful execution
training completed. 0.000/0.000/33.413 (user, system, wall), mem=0 MB
bounds: [0.706041..0.706041] (0.000000) loss=0.616041, iterations=7607
gosdt reported successful execution
training completed. 0.000/0.000/33.448 (user, system, wall), mem=0 MB
bounds: [0.694096..0.694096] (0.000000) loss=0.604096, iterations=7595
gosdt reported successful execution
training completed. 0.000/0.000/34.709 (user, system, wall), mem=0 MB
bounds: [0.692389..0.692389] (0.000000) loss=0.602389, iterations=7569
gosdt reported successful execution
training completed. 0.000/0.000/34.366 (user, system, wall), mem=0 MB
bounds: [0.700920..0.700920] (0.000000) loss=0.580920, iterations=7623
gosdt reported successful execution
training completed. 0.000/0.000/34.103 (user, system, wall), mem=0 MB
bounds: [0.693066..0.693066] (0.000000) loss=0.603066, iterations=7553
here2
gosdt reported successful execution
training completed. 0.000/0.000/32.761 (user, system, wall), mem=0 MB
bounds: [0.756997..0.756997] (0.000000) loss=0.656997, iterations=7607
gosdt reported successful execution
training completed. 0.000/0.000/32.346 (user, system, wall), mem=0 MB
bounds: [0.746758..0.746758] (0.000000) loss=0.646758, iterations=7595
gosdt reported successful execution
training completed. 0.000/0.000/33.433 (user, system, wall), mem=0 MB
bounds: [0.750171..0.750171] (0.000000) loss=0.650171, iterations=7569
gosdt reported successful execution
training completed. 0.000/0.000/32.819 (user, system, wall), mem=0 MB
bounds: [0.752470..0.752470] (0.000000) loss=0.652470, iterations=7623
gosdt reported successful execution
training completed. 0.000/0.000/31.833 (user, system, wall), mem=0 MB
bounds: [0.747359..0.747359] (0.000000) loss=0.647359, iterations=7553

```





here2

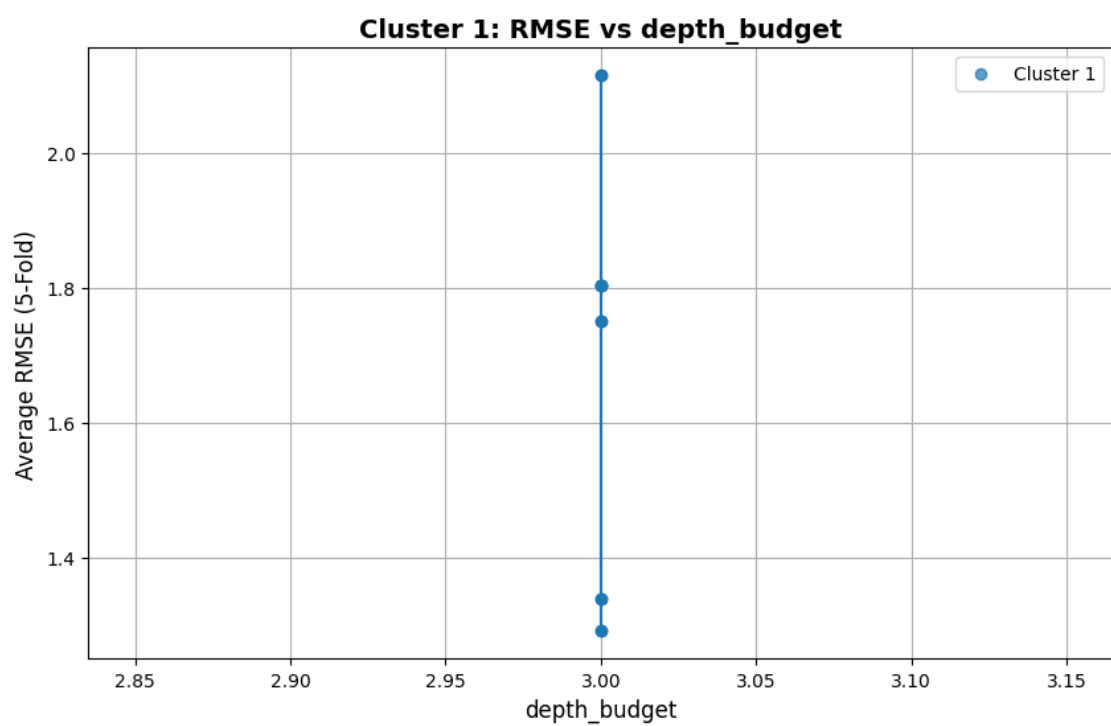
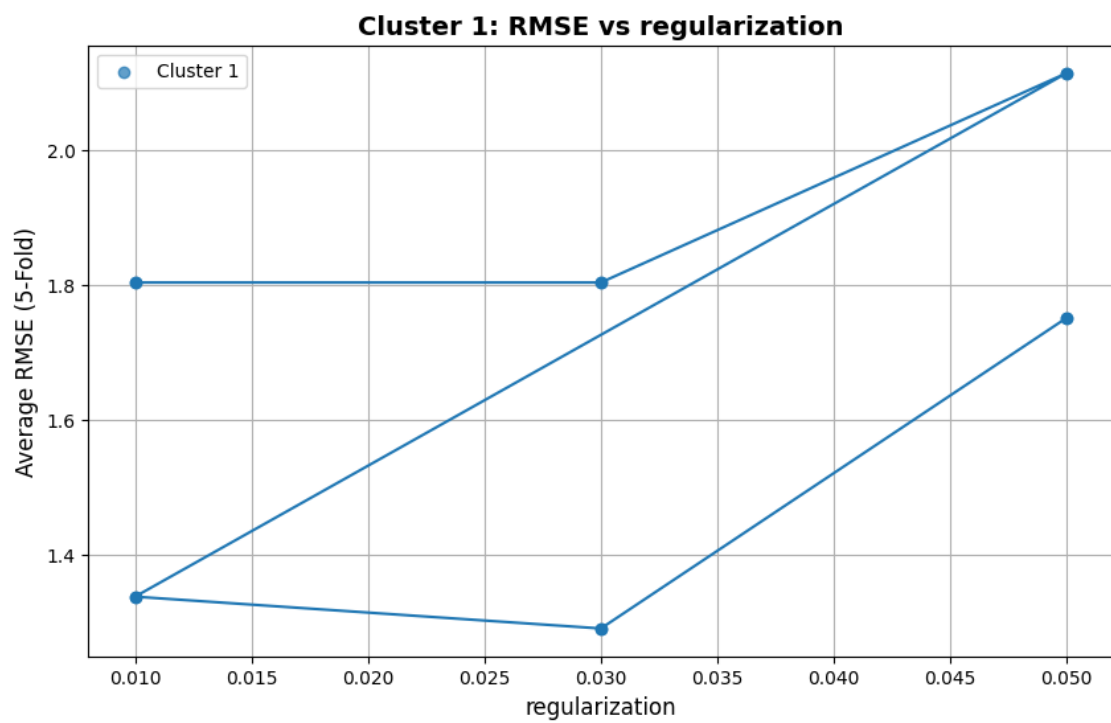


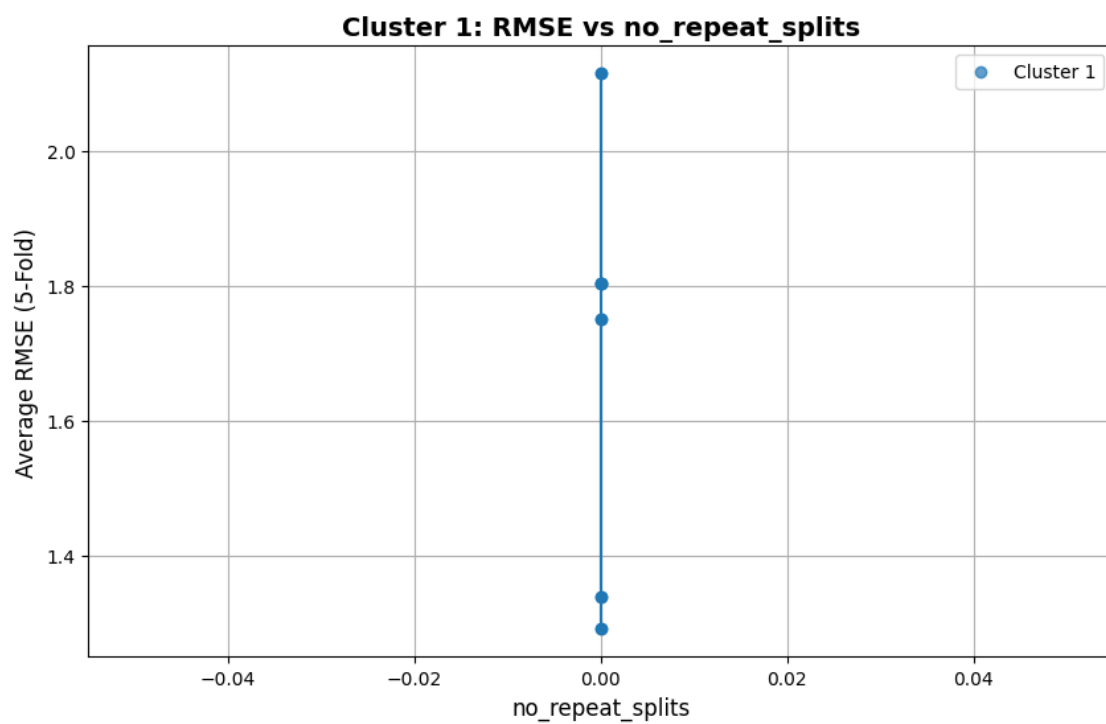
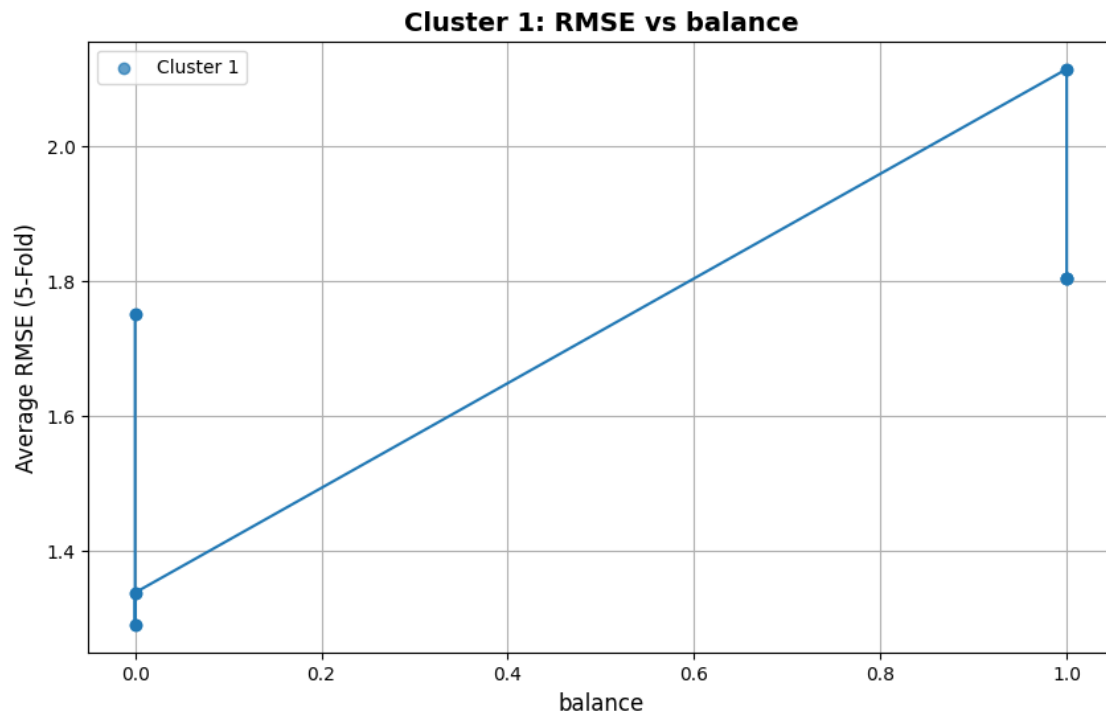
gosdt reported successful execution  
 training completed. 0.000/0.000/46.221 (user, system, wall), mem=0 MB  
 bounds: [0.583086..0.583086] (0.000000) loss=0.543086, iterations=8109  
 gosdt reported successful execution  
 training completed. 0.000/0.000/45.386 (user, system, wall), mem=0 MB  
 bounds: [0.598472..0.598472] (0.000000) loss=0.558472, iterations=8027  
 gosdt reported successful execution  
 training completed. 0.000/0.000/48.076 (user, system, wall), mem=0 MB  
 bounds: [0.592698..0.592698] (0.000000) loss=0.552698, iterations=8131  
 gosdt reported successful execution  
 training completed. 0.000/0.000/46.113 (user, system, wall), mem=0 MB  
 bounds: [0.601907..0.601907] (0.000000) loss=0.561907, iterations=7993  
 gosdt reported successful execution  
 training completed. 0.000/0.000/48.390 (user, system, wall), mem=0 MB  
 bounds: [0.596055..0.596055] (0.000000) loss=0.556055, iterations=8147  
 here2  
 gosdt reported successful execution  
 training completed. 0.000/0.000/46.205 (user, system, wall), mem=0 MB  
 bounds: [0.663086..0.663086] (0.000000) loss=0.543086, iterations=8109  
 gosdt reported successful execution  
 training completed. 0.000/0.000/44.221 (user, system, wall), mem=0 MB  
 bounds: [0.678472..0.678472] (0.000000) loss=0.558472, iterations=8027  
 gosdt reported successful execution  
 training completed. 0.000/0.000/44.862 (user, system, wall), mem=0 MB  
 bounds: [0.672698..0.672698] (0.000000) loss=0.552698, iterations=8131  
 gosdt reported successful execution  
 training completed. 0.000/0.000/43.842 (user, system, wall), mem=0 MB  
 bounds: [0.681907..0.681907] (0.000000) loss=0.561907, iterations=7993  
 gosdt reported successful execution  
 training completed. 0.000/0.000/45.715 (user, system, wall), mem=0 MB  
 bounds: [0.676055..0.676055] (0.000000) loss=0.556055, iterations=8147  
 here2  
 gosdt reported successful execution  
 training completed. 0.000/0.000/43.191 (user, system, wall), mem=0 MB  
 bounds: [0.743086..0.743086] (0.000000) loss=0.543086, iterations=8109  
 gosdt reported successful execution  
 training completed. 0.000/0.000/41.751 (user, system, wall), mem=0 MB  
 bounds: [0.758472..0.758472] (0.000000) loss=0.558472, iterations=8027  
 gosdt reported successful execution  
 training completed. 0.000/0.000/43.428 (user, system, wall), mem=0 MB  
 bounds: [0.752698..0.752698] (0.000000) loss=0.552698, iterations=8131  
 gosdt reported successful execution  
 training completed. 0.000/0.000/41.830 (user, system, wall), mem=0 MB  
 bounds: [0.756155..0.756155] (0.000000) loss=0.606155, iterations=7993  
 gosdt reported successful execution  
 training completed. 0.000/0.000/43.287 (user, system, wall), mem=0 MB  
 bounds: [0.754341..0.754341] (0.000000) loss=0.604341, iterations=8147  
 here2

```

gosdt reported successful execution
training completed. 0.000/0.000/46.312 (user, system, wall), mem=0 MB
bounds: [0.533902..0.533902] (0.000000) loss=0.493902, iterations=8109
gosdt reported successful execution
training completed. 0.000/0.000/46.653 (user, system, wall), mem=0 MB
bounds: [0.562533..0.562533] (0.000000) loss=0.522534, iterations=8027
gosdt reported successful execution
training completed. 0.000/0.000/47.704 (user, system, wall), mem=0 MB
bounds: [0.560097..0.560097] (0.000000) loss=0.520097, iterations=8131
gosdt reported successful execution
training completed. 0.000/0.000/45.208 (user, system, wall), mem=0 MB
bounds: [0.558879..0.558879] (0.000000) loss=0.518879, iterations=7993
gosdt reported successful execution
training completed. 0.000/0.000/47.586 (user, system, wall), mem=0 MB
bounds: [0.558879..0.558879] (0.000000) loss=0.518879, iterations=8147
here2
gosdt reported successful execution
training completed. 0.000/0.000/43.993 (user, system, wall), mem=0 MB
bounds: [0.610732..0.610732] (0.000000) loss=0.520732, iterations=8109
gosdt reported successful execution
training completed. 0.000/0.000/43.025 (user, system, wall), mem=0 MB
bounds: [0.634458..0.634458] (0.000000) loss=0.544458, iterations=8027
gosdt reported successful execution
training completed. 0.000/0.000/44.265 (user, system, wall), mem=0 MB
bounds: [0.635676..0.635676] (0.000000) loss=0.545676, iterations=8131
gosdt reported successful execution
training completed. 0.000/0.000/42.605 (user, system, wall), mem=0 MB
bounds: [0.633240..0.633240] (0.000000) loss=0.543240, iterations=7993
gosdt reported successful execution
training completed. 0.000/0.000/44.277 (user, system, wall), mem=0 MB
bounds: [0.635676..0.635676] (0.000000) loss=0.545676, iterations=8147
here2
gosdt reported successful execution
training completed. 0.000/0.000/42.577 (user, system, wall), mem=0 MB
bounds: [0.670732..0.670732] (0.000000) loss=0.520732, iterations=8109
gosdt reported successful execution
training completed. 0.000/0.000/42.193 (user, system, wall), mem=0 MB
bounds: [0.676066..0.676066] (0.000000) loss=0.626066, iterations=8027
gosdt reported successful execution
training completed. 0.000/0.000/42.293 (user, system, wall), mem=0 MB
bounds: [0.694336..0.694336] (0.000000) loss=0.644336, iterations=8131
gosdt reported successful execution
training completed. 0.000/0.000/41.830 (user, system, wall), mem=0 MB
bounds: [0.666322..0.666322] (0.000000) loss=0.616322, iterations=7993
gosdt reported successful execution
training completed. 0.000/0.000/43.952 (user, system, wall), mem=0 MB
bounds: [0.695676..0.695676] (0.000000) loss=0.545676, iterations=8147

```





Trained models and their RMSEs:

Cluster 7 - Best RMSE: 1.2486849347337419, Params: {'regularization': 0.05, 'depth\_budget': 3, 'balance': False, 'no\_repeat\_splits': False}  
 Cluster 6 - Best RMSE: 1.2136896934008874, Params: {'regularization': 0.05, 'depth\_budget': 3, 'balance': False, 'no\_repeat\_splits': False}  
 Cluster 0 - Best RMSE: 1.4188985118640591, Params: {'regularization': 0.05, 'depth\_budget': 3, 'balance': False, 'no\_repeat\_splits': False}  
 Cluster 5 - Best RMSE: 1.3552193970356783, Params: {'regularization': 0.05, 'depth\_budget': 3, 'balance': False, 'no\_repeat\_splits': False}  
 Cluster 4 - Best RMSE: 1.4526707545078668, Params: {'regularization': 0.05, 'depth\_budget': 3, 'balance': False, 'no\_repeat\_splits': False}  
 Cluster 3 - Best RMSE: 1.2561241614262832, Params: {'regularization': 0.05, 'depth\_budget': 3, 'balance': False, 'no\_repeat\_splits': False}  
 Cluster 8 - Best RMSE: 1.2076106969237628, Params: {'regularization': 0.05, 'depth\_budget': 3, 'balance': False, 'no\_repeat\_splits': False}  
 Cluster 2 - Best RMSE: 1.507358000123405, Params: {'regularization': 0.05, 'depth\_budget': 3, 'balance': False, 'no\_repeat\_splits': False}  
 Cluster 9 - Best RMSE: 1.1336552827262887, Params: {'regularization': 0.05, 'depth\_budget': 3, 'balance': False, 'no\_repeat\_splits': False}  
 Cluster 1 - Best RMSE: 1.2910291896063077, Params: {'regularization': 0.05, 'depth\_budget': 3, 'balance': False, 'no\_repeat\_splits': False}

### 3 Plot RMSE against cluster size

```
[ ]: from sklearn.linear_model import LinearRegression

# Data
clusters = ["Cluster 7", "Cluster 6", "Cluster 0", "Cluster 5", "Cluster 4",
            "Cluster 3", "Cluster 8", "Cluster 2", "Cluster 9", "Cluster 1"]
properties = [2407, 1917, 2037, 1212, 2630, 1791, 588, 1355, 733, 1026]
rmsees = [1.2486849347337419, 1.2136896934008874, 1.4188985118640591,
          1.3552193970356783, 1.4526707545078668, 1.2561241614262832,
          1.2076106969237628, 1.507358000123405, 1.1336552827262887,
          1.2910291896063077]

# Reshape data for linear regression
X = np.array(properties).reshape(-1, 1)
y = np.array(rmsees)

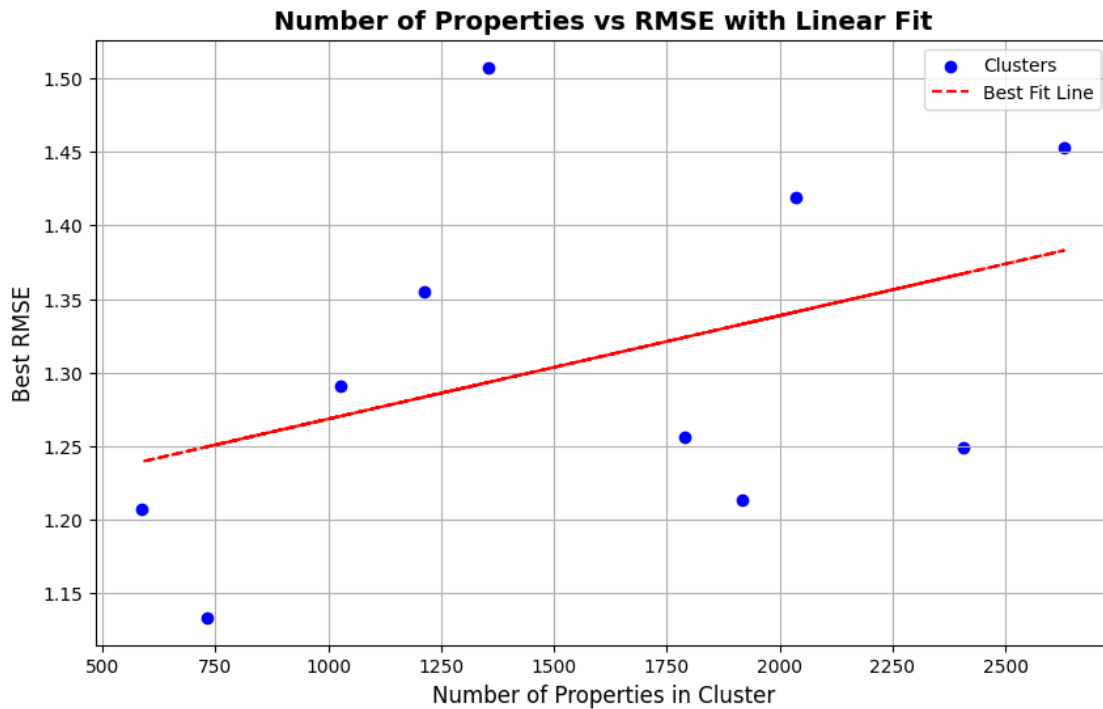
# Fit a linear regression model
model = LinearRegression()
model.fit(X, y)
y_pred = model.predict(X)

# Plotting
plt.figure(figsize=(10, 6))
plt.scatter(properties, rmsees, color='blue', label='Clusters')
plt.plot(properties, y_pred, color='red', linestyle='--', label='Best Fit Line')
```

```

# Add labels, title, and legend
plt.xlabel("Number of Properties in Cluster", fontsize=12)
plt.ylabel("Best RMSE", fontsize=12)
plt.title("Number of Properties vs RMSE with Linear Fit", fontsize=14,
          fontweight='bold')
plt.legend()
plt.grid(True)
plt.show()

```



### 3.1 Predict on the Test Dataset

```

[ ]: submission_data = []

# Loop through each test sample
for idx, test_sample in test_df.iterrows():
    # Drop 'id' and 'cluster' columns for prediction
    prediction_features = test_sample.drop(['id', 'cluster'])

    # Identify the cluster
    test_cluster_id = int(test_sample['cluster']) # Get the cluster ID for the
    # test sample

    # Retrieve the model

```

```

    model_info = next((item for item in models if item['cluster'] ==
↳test_cluster_id), None)

    if model_info:
        model = model_info['model'] # Extract the trained model
        # Make the prediction (reshape the features to be 2D, as the model
↳expects it)
        prediction = model.predict([prediction_features])[0] # Ensure correct
↳shape for prediction
    else:
        prediction = None # If no model exists for this cluster

    # Append the 'id' and the prediction to the submission data
    submission_data.append([test_sample['id'], prediction])

# Convert the submission data into a DataFrame and save it
submission_df = pd.DataFrame(submission_data, columns=['id', 'price'])
submission_df.to_csv('submission.csv', index=False)

```