KNN_Regression

November 27, 2025

```python
[1]: # -*- coding: utf-8 -*-
"""
Stage 1 - GPU Accelerated KNN Regression using PyTorch
"""
"""
Stage 1 - KNN Regression (from scratch)
Performs:
- 80/20 split
- 70/30 split
- Euclidean / Manhattan / Minkowski distances
- Cross-validation
- Grid search for best k
- Linear Regression baseline
- Saves results to CSV
"""

import numpy as np
import pandas as pd
import torch
import csv

from sklearn.model_selection import train_test_split, KFold
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import mean_squared_error, r2_score
from sklearn.linear_model import LinearRegression
from ucimlrepo import fetch_ucirepo


# ---------------- GPU DEVICE ----------------
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
print("Using device:", device)


# ---------------- Distance Functions ----------------
def compute_distances(x, X_train, metric="euclidean", p=3):
    if metric == "euclidean":
        return torch.norm(X_train - x, dim=1)
```

```python
    elif metric == "manhattan":
        return torch.sum(torch.abs(X_train - x), dim=1)

    elif metric == "minkowski":
        return torch.sum(torch.abs(X_train - x)**p, dim=1) ** (1/p)


# ----------------- KNN Regressor -----------------
class KNNRegressorTorch:
    def __init__(self, k=5, distance="euclidean", weights="uniform", p=3):
        self.k = k
        self.distance = distance
        self.weights = weights
        self.p = p

    def fit(self, X, y):
        print("   Fit: KNN (k={}, distance={})".format(self.k, self.distance))
        self.X_train = torch.tensor(X, dtype=torch.float32).to(device)
        self.y_train = torch.tensor(y.values, dtype=torch.float32).to(device)

    def predict(self, X):
        print("   Predict: Using KNN with distance =", self.distance)
        X = torch.tensor(X, dtype=torch.float32).to(device)
        preds = []

        for idx, x in enumerate(X):
            if idx % 100 == 0:
                print("      Predicted", idx, "samples")

            dists = compute_distances(x, self.X_train, self.distance, self.p)
            knn_idx = torch.topk(dists, self.k, largest=False).indices
            knn_vals = self.y_train[knn_idx]

            if self.weights == "uniform":
                pred = knn_vals.mean()

            else:
                eps = 1e-8
                weights = 1 / (dists[knn_idx] + eps)
                pred = (weights * knn_vals).sum() / weights.sum()

            preds.append(pred.item())

        return np.array(preds)
```

```python
# ---------------- Cross Validation ----------------
def cross_val_score_knn_torch(X, y, k_values, distance, p=3):
    results = {}
    kf = KFold(n_splits=5, shuffle=True, random_state=42)

    X_np = X
    y_np = y.values

    print("Grid Search for distance:", distance)

    for k in k_values:
        print("  Testing k =", k)

        mse_list = []
        fold_num = 1

        for train_idx, val_idx in kf.split(X_np):
            print("    Fold", fold_num, "of 5")

            model = KNNRegressorTorch(k=k, distance=distance, p=p)
            model.fit(X_np[train_idx], y.iloc[train_idx])

            y_pred = model.predict(X_np[val_idx])
            mse_list.append(mean_squared_error(y_np[val_idx], y_pred))

            fold_num += 1

        avg_mse = np.mean(mse_list)
        print("    Average CV MSE for k =", k, "is", avg_mse)

        results[k] = avg_mse

    print("Finished Grid Search for", distance)
    return results


# ---------------- Experiment Function ----------------
def run_experiment_torch(X_scaled, y, test_size, k_grid):

    print("Running experiment with test size =", test_size)

    X_train, X_test, y_train, y_test = train_test_split(
        X_scaled, y, test_size=test_size, random_state=42
    )

    # Euclidean
    print("\nDistance: Euclidean")
```

```python
cv_eu = cross_val_score_knn_torch(X_scaled, y, k_grid, "euclidean")
best_k_eu = min(cv_eu, key=cv_eu.get)
print("Best k (Euclidean):", best_k_eu)

model = KNNRegressorTorch(k=best_k_eu, distance="euclidean")
model.fit(X_train, y_train)
pred_eu = model.predict(X_test)
mse_eu = mean_squared_error(y_test, pred_eu)
rmse_eu = np.sqrt(mse_eu)
r2_eu = r2_score(y_test, pred_eu)

# Manhattan
print("\nDistance: Manhattan")
cv_manh = cross_val_score_knn_torch(X_scaled, y, k_grid, "manhattan")
best_k_manh = min(cv_manh, key=cv_manh.get)
print("Best k (Manhattan):", best_k_manh)

model = KNNRegressorTorch(k=best_k_manh, distance="manhattan")
model.fit(X_train, y_train)
pred_manh = model.predict(X_test)
mse_manh = mean_squared_error(y_test, pred_manh)
rmse_manh = np.sqrt(mse_manh)
r2_manh = r2_score(y_test, pred_manh)

# Minkowski
print("\nDistance: Minkowski (p=3)")
cv_mink = cross_val_score_knn_torch(X_scaled, y, k_grid, "minkowski", p=3)
best_k_mink = min(cv_mink, key=cv_mink.get)
print("Best k (Minkowski):", best_k_mink)

model = KNNRegressorTorch(k=best_k_mink, distance="minkowski", p=3)
model.fit(X_train, y_train)
pred_mink = model.predict(X_test)
mse_mink = mean_squared_error(y_test, pred_mink)
rmse_mink = np.sqrt(mse_mink)
r2_mink = r2_score(y_test, pred_mink)

# Baseline Linear Regression
print("\nRunning Linear Regression baseline")
lr = LinearRegression()
lr.fit(X_train, y_train)
lr_pred = lr.predict(X_test)
mse_lr = mean_squared_error(y_test, lr_pred)
rmse_lr = np.sqrt(mse_lr)
r2_lr = r2_score(y_test, lr_pred)
print("Baseline completed")
```

```python
    return {
        "euclidean": (best_k_eu, mse_eu, rmse_eu, r2_eu),
        "manhattan": (best_k_manh, mse_manh, rmse_manh, r2_manh),
        "minkowski": (best_k_mink, mse_mink, rmse_mink, r2_mink),
        "baseline": (mse_lr, rmse_lr, r2_lr)
    }



# ---------------- MAIN SCRIPT ----------------

print("Loading Wine Quality Dataset...")
wine_quality = fetch_ucirepo(id=186)
X = wine_quality.data.features
y = wine_quality.data.targets
print("Dataset loaded.")

print("Scaling features...")
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)
print("Scaling completed.")

k_grid = [1, 3, 5, 7, 9, 11, 15, 19]

results_80 = run_experiment_torch(X_scaled, y, test_size=0.2, k_grid=k_grid)
results_70 = run_experiment_torch(X_scaled, y, test_size=0.3, k_grid=k_grid)

print("Saving results to CSV file: stage1_results.csv")

with open("stage1_results.csv", "w", newline="") as f:
    writer = csv.writer(f)
    writer.writerow(["Split", "Model", "Distance", "Best_k", "MSE", "RMSE",␣
 ↪"R2"])

    writer.writerow(["80/20", "KNN", "Euclidean", *results_80["euclidean"]])
    writer.writerow(["80/20", "KNN", "Manhattan", *results_80["manhattan"]])
    writer.writerow(["80/20", "KNN", "Minkowski", *results_80["minkowski"]])
    writer.writerow(["80/20", "Linear Regression", "NA", "NA",␣
 ↪*results_80["baseline"]])

    writer.writerow(["70/30", "KNN", "Euclidean", *results_70["euclidean"]])
    writer.writerow(["70/30", "KNN", "Manhattan", *results_70["manhattan"]])
    writer.writerow(["70/30", "KNN", "Minkowski", *results_70["minkowski"]])
    writer.writerow(["70/30", "Linear Regression", "NA", "NA",␣
 ↪*results_70["baseline"]])

print("Done.")
```

```
Using device: cuda
Loading Wine Quality Dataset…
Dataset loaded.
Scaling features…
Scaling completed.
Running experiment with test size = 0.2

Distance: Euclidean
Grid Search for distance: euclidean
  Testing k = 1
      Fold 1 of 5
    Fit: KNN (k=1, distance=euclidean)
    Predict: Using KNN with distance = euclidean
        Predicted 0 samples
        Predicted 100 samples
        Predicted 200 samples
        Predicted 300 samples
        Predicted 400 samples
        Predicted 500 samples
        Predicted 600 samples
        Predicted 700 samples
        Predicted 800 samples
        Predicted 900 samples
        Predicted 1000 samples
        Predicted 1100 samples
        Predicted 1200 samples
      Fold 2 of 5
    Fit: KNN (k=1, distance=euclidean)
    Predict: Using KNN with distance = euclidean
        Predicted 0 samples
        Predicted 100 samples
        Predicted 200 samples
        Predicted 300 samples
        Predicted 400 samples
        Predicted 500 samples
        Predicted 600 samples
        Predicted 700 samples
        Predicted 800 samples
        Predicted 900 samples
        Predicted 1000 samples
        Predicted 1100 samples
        Predicted 1200 samples
      Fold 3 of 5
    Fit: KNN (k=1, distance=euclidean)
    Predict: Using KNN with distance = euclidean
        Predicted 0 samples
        Predicted 100 samples
        Predicted 200 samples
```

```
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Fold 4 of 5
 Fit: KNN (k=1, distance=euclidean)
 Predict: Using KNN with distance = euclidean
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Fold 5 of 5
 Fit: KNN (k=1, distance=euclidean)
 Predict: Using KNN with distance = euclidean
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Average CV MSE for k = 1 is 0.6181254219221887
 Testing k = 3
    Fold 1 of 5
 Fit: KNN (k=3, distance=euclidean)
 Predict: Using KNN with distance = euclidean
      Predicted 0 samples
```

```
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 2 of 5
Fit: KNN (k=3, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=3, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=3, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
```

```
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 5 of 5
 Fit: KNN (k=3, distance=euclidean)
 Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
   Average CV MSE for k = 3 is 0.5147669963160465
 Testing k = 5
   Fold 1 of 5
 Fit: KNN (k=5, distance=euclidean)
 Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 2 of 5
 Fit: KNN (k=5, distance=euclidean)
```

```
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=5, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=5, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 5 of 5
Fit: KNN (k=5, distance=euclidean)
```

```
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
   Average CV MSE for k = 5 is 0.4902327512557251
Testing k = 7
   Fold 1 of 5
 Fit: KNN (k=7, distance=euclidean)
 Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
   Fold 2 of 5
 Fit: KNN (k=7, distance=euclidean)
 Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
```

```
  Fold 3 of 5
Fit: KNN (k=7, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=7, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 5 of 5
Fit: KNN (k=7, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
```

```
   Average CV MSE for k = 7 is 0.4826653993245579
Testing k = 9
   Fold 1 of 5
 Fit: KNN (k=9, distance=euclidean)
 Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Fold 2 of 5
 Fit: KNN (k=9, distance=euclidean)
 Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Fold 3 of 5
 Fit: KNN (k=9, distance=euclidean)
 Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
```

```
      Predicted 1100 samples
      Predicted 1200 samples
    Fold 4 of 5
 Fit: KNN (k=9, distance=euclidean)
 Predict: Using KNN with distance = euclidean
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Fold 5 of 5
 Fit: KNN (k=9, distance=euclidean)
 Predict: Using KNN with distance = euclidean
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Average CV MSE for k = 9 is 0.4842311761775517
 Testing k = 11
    Fold 1 of 5
 Fit: KNN (k=11, distance=euclidean)
 Predict: Using KNN with distance = euclidean
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
```

```
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 2 of 5
Fit: KNN (k=11, distance=euclidean)
Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=11, distance=euclidean)
Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=11, distance=euclidean)
Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
```

```
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 5 of 5
Fit: KNN (k=11, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Average CV MSE for k = 11 is 0.48664736144092763
Testing k = 15
  Fold 1 of 5
Fit: KNN (k=15, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 2 of 5
Fit: KNN (k=15, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
```

```
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=15, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=15, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 5 of 5
Fit: KNN (k=15, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
```

```
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Average CV MSE for k = 15 is 0.4887360075411709
 Testing k = 19
   Fold 1 of 5
 Fit: KNN (k=19, distance=euclidean)
 Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Fold 2 of 5
 Fit: KNN (k=19, distance=euclidean)
 Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Fold 3 of 5
 Fit: KNN (k=19, distance=euclidean)
 Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
```

```
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Fold 4 of 5
  Fit: KNN (k=19, distance=euclidean)
  Predict: Using KNN with distance = euclidean
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Fold 5 of 5
  Fit: KNN (k=19, distance=euclidean)
  Predict: Using KNN with distance = euclidean
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Average CV MSE for k = 19 is 0.49255148928453796
Finished Grid Search for euclidean
Best k (Euclidean): 7
  Fit: KNN (k=7, distance=euclidean)
  Predict: Using KNN with distance = euclidean
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
```

```
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples


Distance: Manhattan
Grid Search for distance: manhattan
  Testing k = 1
     Fold 1 of 5
   Fit: KNN (k=1, distance=manhattan)
   Predict: Using KNN with distance = manhattan
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
     Fold 2 of 5
   Fit: KNN (k=1, distance=manhattan)
   Predict: Using KNN with distance = manhattan
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
     Fold 3 of 5
   Fit: KNN (k=1, distance=manhattan)
```

```
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
 Fold 4 of 5
Fit: KNN (k=1, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
 Fold 5 of 5
Fit: KNN (k=1, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
 Average CV MSE for k = 1 is 0.60012115828744
Testing k = 3
```

```
  Fold 1 of 5
Fit: KNN (k=3, distance=manhattan)
Predict: Using KNN with distance = manhattan
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
  Fold 2 of 5
Fit: KNN (k=3, distance=manhattan)
Predict: Using KNN with distance = manhattan
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=3, distance=manhattan)
Predict: Using KNN with distance = manhattan
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
```

```
   Fold 4 of 5
 Fit: KNN (k=3, distance=manhattan)
 Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Fold 5 of 5
 Fit: KNN (k=3, distance=manhattan)
 Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Average CV MSE for k = 3 is 0.5086804823745356
Testing k = 5
   Fold 1 of 5
 Fit: KNN (k=5, distance=manhattan)
 Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
```

```
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 2 of 5
Fit: KNN (k=5, distance=manhattan)
Predict: Using KNN with distance = manhattan
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 3 of 5
Fit: KNN (k=5, distance=manhattan)
Predict: Using KNN with distance = manhattan
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 4 of 5
Fit: KNN (k=5, distance=manhattan)
Predict: Using KNN with distance = manhattan
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
```

```
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 5 of 5
 Fit: KNN (k=5, distance=manhattan)
 Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
   Average CV MSE for k = 5 is 0.4882050392291152
Testing k = 7
  Fold 1 of 5
 Fit: KNN (k=7, distance=manhattan)
 Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 2 of 5
 Fit: KNN (k=7, distance=manhattan)
 Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
```

```
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=7, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=7, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 5 of 5
Fit: KNN (k=7, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
```

```
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Average CV MSE for k = 7 is 0.4863480978659928
 Testing k = 9
   Fold 1 of 5
 Fit: KNN (k=9, distance=manhattan)
 Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Fold 2 of 5
 Fit: KNN (k=9, distance=manhattan)
 Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Fold 3 of 5
 Fit: KNN (k=9, distance=manhattan)
 Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
```

```
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 4 of 5
 Fit: KNN (k=9, distance=manhattan)
 Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 5 of 5
 Fit: KNN (k=9, distance=manhattan)
 Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Average CV MSE for k = 9 is 0.4837643886862325
Testing k = 11
  Fold 1 of 5
 Fit: KNN (k=11, distance=manhattan)
 Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
```

```
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 2 of 5
Fit: KNN (k=11, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=11, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=11, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
```

```
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Fold 5 of 5
 Fit: KNN (k=11, distance=manhattan)
 Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Average CV MSE for k = 11 is 0.4809257289504811
 Testing k = 15
   Fold 1 of 5
 Fit: KNN (k=15, distance=manhattan)
 Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Fold 2 of 5
 Fit: KNN (k=15, distance=manhattan)
 Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
```

```
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=15, distance=manhattan)
Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=15, distance=manhattan)
Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 5 of 5
Fit: KNN (k=15, distance=manhattan)
Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
```

```
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Average CV MSE for k = 15 is 0.4856087159121366
 Testing k = 19
   Fold 1 of 5
 Fit: KNN (k=19, distance=manhattan)
 Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
   Fold 2 of 5
 Fit: KNN (k=19, distance=manhattan)
 Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
   Fold 3 of 5
 Fit: KNN (k=19, distance=manhattan)
 Predict: Using KNN with distance = manhattan
   Predicted 0 samples
```

```
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Fold 4 of 5
  Fit: KNN (k=19, distance=manhattan)
  Predict: Using KNN with distance = manhattan
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Fold 5 of 5
  Fit: KNN (k=19, distance=manhattan)
  Predict: Using KNN with distance = manhattan
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Average CV MSE for k = 19 is 0.48717413211159366
Finished Grid Search for manhattan
Best k (Manhattan): 11
  Fit: KNN (k=11, distance=manhattan)
```

```
   Predict: Using KNN with distance = manhattan
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples


Distance: Minkowski (p=3)
Grid Search for distance: minkowski
  Testing k = 1
     Fold 1 of 5
   Fit: KNN (k=1, distance=minkowski)
   Predict: Using KNN with distance = minkowski
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
     Fold 2 of 5
   Fit: KNN (k=1, distance=minkowski)
   Predict: Using KNN with distance = minkowski
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
```

```
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 3 of 5
Fit: KNN (k=1, distance=minkowski)
Predict: Using KNN with distance = minkowski
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 4 of 5
Fit: KNN (k=1, distance=minkowski)
Predict: Using KNN with distance = minkowski
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 5 of 5
Fit: KNN (k=1, distance=minkowski)
Predict: Using KNN with distance = minkowski
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
```

```
   Predicted 1100 samples
   Predicted 1200 samples
  Average CV MSE for k = 1 is 0.6316682655296975
Testing k = 3
  Fold 1 of 5
Fit: KNN (k=3, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 2 of 5
Fit: KNN (k=3, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=3, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
```

```
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=3, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 5 of 5
Fit: KNN (k=3, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Average CV MSE for k = 3 is 0.5154682173084904
Testing k = 5
  Fold 1 of 5
Fit: KNN (k=5, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
```

```
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 2 of 5
Fit: KNN (k=5, distance=minkowski)
Predict: Using KNN with distance = minkowski
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=5, distance=minkowski)
Predict: Using KNN with distance = minkowski
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=5, distance=minkowski)
Predict: Using KNN with distance = minkowski
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
```

```
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 5 of 5
 Fit: KNN (k=5, distance=minkowski)
 Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
   Average CV MSE for k = 5 is 0.49037387319057296
Testing k = 7
  Fold 1 of 5
 Fit: KNN (k=7, distance=minkowski)
 Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 2 of 5
 Fit: KNN (k=7, distance=minkowski)
 Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
```

```
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 3 of 5
Fit: KNN (k=7, distance=minkowski)
Predict: Using KNN with distance = minkowski
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 4 of 5
Fit: KNN (k=7, distance=minkowski)
Predict: Using KNN with distance = minkowski
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 5 of 5
Fit: KNN (k=7, distance=minkowski)
Predict: Using KNN with distance = minkowski
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
```

```
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
    Average CV MSE for k = 7 is 0.4844860028733596
Testing k = 9
    Fold 1 of 5
 Fit: KNN (k=9, distance=minkowski)
 Predict: Using KNN with distance = minkowski
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
    Fold 2 of 5
 Fit: KNN (k=9, distance=minkowski)
 Predict: Using KNN with distance = minkowski
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
    Fold 3 of 5
 Fit: KNN (k=9, distance=minkowski)
 Predict: Using KNN with distance = minkowski
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
```

```
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 4 of 5
 Fit: KNN (k=9, distance=minkowski)
 Predict: Using KNN with distance = minkowski
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 5 of 5
 Fit: KNN (k=9, distance=minkowski)
 Predict: Using KNN with distance = minkowski
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Average CV MSE for k = 9 is 0.4875670170853498
 Testing k = 11
  Fold 1 of 5
 Fit: KNN (k=11, distance=minkowski)
 Predict: Using KNN with distance = minkowski
    Predicted 0 samples
```

```
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 2 of 5
Fit: KNN (k=11, distance=minkowski)
Predict: Using KNN with distance = minkowski
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=11, distance=minkowski)
Predict: Using KNN with distance = minkowski
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=11, distance=minkowski)
Predict: Using KNN with distance = minkowski
    Predicted 0 samples
```

```
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 5 of 5
 Fit: KNN (k=11, distance=minkowski)
 Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
   Average CV MSE for k = 11 is 0.49063564537851956
Testing k = 15
  Fold 1 of 5
 Fit: KNN (k=15, distance=minkowski)
 Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 2 of 5
 Fit: KNN (k=15, distance=minkowski)
```

```
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=15, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=15, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 5 of 5
Fit: KNN (k=15, distance=minkowski)
```

```
  Predict: Using KNN with distance = minkowski
     Predicted 0 samples
     Predicted 100 samples
     Predicted 200 samples
     Predicted 300 samples
     Predicted 400 samples
     Predicted 500 samples
     Predicted 600 samples
     Predicted 700 samples
     Predicted 800 samples
     Predicted 900 samples
     Predicted 1000 samples
     Predicted 1100 samples
     Predicted 1200 samples
     Average CV MSE for k = 15 is 0.49593958869960586
Testing k = 19
     Fold 1 of 5
 Fit: KNN (k=19, distance=minkowski)
 Predict: Using KNN with distance = minkowski
     Predicted 0 samples
     Predicted 100 samples
     Predicted 200 samples
     Predicted 300 samples
     Predicted 400 samples
     Predicted 500 samples
     Predicted 600 samples
     Predicted 700 samples
     Predicted 800 samples
     Predicted 900 samples
     Predicted 1000 samples
     Predicted 1100 samples
     Predicted 1200 samples
     Fold 2 of 5
 Fit: KNN (k=19, distance=minkowski)
 Predict: Using KNN with distance = minkowski
     Predicted 0 samples
     Predicted 100 samples
     Predicted 200 samples
     Predicted 300 samples
     Predicted 400 samples
     Predicted 500 samples
     Predicted 600 samples
     Predicted 700 samples
     Predicted 800 samples
     Predicted 900 samples
     Predicted 1000 samples
     Predicted 1100 samples
     Predicted 1200 samples
```

```
  Fold 3 of 5
Fit: KNN (k=19, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=19, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 5 of 5
Fit: KNN (k=19, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
```

```
      Average CV MSE for k = 19 is 0.4993375809582264
Finished Grid Search for minkowski
Best k (Minkowski): 7
   Fit: KNN (k=7, distance=minkowski)
   Predict: Using KNN with distance = minkowski
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples


Running Linear Regression baseline
Baseline completed
Running experiment with test size = 0.3

Distance: Euclidean
Grid Search for distance: euclidean
  Testing k = 1
     Fold 1 of 5
   Fit: KNN (k=1, distance=euclidean)
   Predict: Using KNN with distance = euclidean
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
     Fold 2 of 5
   Fit: KNN (k=1, distance=euclidean)
   Predict: Using KNN with distance = euclidean
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
```

```
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=1, distance=euclidean)
Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=1, distance=euclidean)
Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 5 of 5
Fit: KNN (k=1, distance=euclidean)
Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
```

```
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Average CV MSE for k = 1 is 0.6181254219221887
Testing k = 3
  Fold 1 of 5
 Fit: KNN (k=3, distance=euclidean)
 Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 2 of 5
 Fit: KNN (k=3, distance=euclidean)
 Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 3 of 5
 Fit: KNN (k=3, distance=euclidean)
 Predict: Using KNN with distance = euclidean
    Predicted 0 samples
```

```
        Predicted 100 samples
        Predicted 200 samples
        Predicted 300 samples
        Predicted 400 samples
        Predicted 500 samples
        Predicted 600 samples
        Predicted 700 samples
        Predicted 800 samples
        Predicted 900 samples
        Predicted 1000 samples
        Predicted 1100 samples
        Predicted 1200 samples
      Fold 4 of 5
 Fit: KNN (k=3, distance=euclidean)
 Predict: Using KNN with distance = euclidean
        Predicted 0 samples
        Predicted 100 samples
        Predicted 200 samples
        Predicted 300 samples
        Predicted 400 samples
        Predicted 500 samples
        Predicted 600 samples
        Predicted 700 samples
        Predicted 800 samples
        Predicted 900 samples
        Predicted 1000 samples
        Predicted 1100 samples
        Predicted 1200 samples
      Fold 5 of 5
 Fit: KNN (k=3, distance=euclidean)
 Predict: Using KNN with distance = euclidean
        Predicted 0 samples
        Predicted 100 samples
        Predicted 200 samples
        Predicted 300 samples
        Predicted 400 samples
        Predicted 500 samples
        Predicted 600 samples
        Predicted 700 samples
        Predicted 800 samples
        Predicted 900 samples
        Predicted 1000 samples
        Predicted 1100 samples
        Predicted 1200 samples
      Average CV MSE for k = 3 is 0.5147669963160465
 Testing k = 5
      Fold 1 of 5
 Fit: KNN (k=5, distance=euclidean)
```

```
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 2 of 5
Fit: KNN (k=5, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=5, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=5, distance=euclidean)
```

```
Predict: Using KNN with distance = euclidean
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 5 of 5
Fit: KNN (k=5, distance=euclidean)
Predict: Using KNN with distance = euclidean
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Average CV MSE for k = 5 is 0.4902327512557251
Testing k = 7
 Fold 1 of 5
Fit: KNN (k=7, distance=euclidean)
Predict: Using KNN with distance = euclidean
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
```

```
  Fold 2 of 5
Fit: KNN (k=7, distance=euclidean)
Predict: Using KNN with distance = euclidean
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=7, distance=euclidean)
Predict: Using KNN with distance = euclidean
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=7, distance=euclidean)
Predict: Using KNN with distance = euclidean
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
```

```
  Fold 5 of 5
 Fit: KNN (k=7, distance=euclidean)
 Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Average CV MSE for k = 7 is 0.4826653993245579
Testing k = 9
   Fold 1 of 5
 Fit: KNN (k=9, distance=euclidean)
 Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Fold 2 of 5
 Fit: KNN (k=9, distance=euclidean)
 Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
```

```
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 3 of 5
Fit: KNN (k=9, distance=euclidean)
Predict: Using KNN with distance = euclidean
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 4 of 5
Fit: KNN (k=9, distance=euclidean)
Predict: Using KNN with distance = euclidean
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 5 of 5
Fit: KNN (k=9, distance=euclidean)
Predict: Using KNN with distance = euclidean
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
```

```
   Predicted 1100 samples
   Predicted 1200 samples
  Average CV MSE for k = 9 is 0.4842311761775517
Testing k = 11
   Fold 1 of 5
 Fit: KNN (k=11, distance=euclidean)
 Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
   Fold 2 of 5
 Fit: KNN (k=11, distance=euclidean)
 Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
   Fold 3 of 5
 Fit: KNN (k=11, distance=euclidean)
 Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
```

```
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=11, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 5 of 5
Fit: KNN (k=11, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Average CV MSE for k = 11 is 0.48664736144092763
Testing k = 15
  Fold 1 of 5
Fit: KNN (k=15, distance=euclidean)
Predict: Using KNN with distance = euclidean
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
```

```
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 2 of 5
Fit: KNN (k=15, distance=euclidean)
Predict: Using KNN with distance = euclidean
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 3 of 5
Fit: KNN (k=15, distance=euclidean)
Predict: Using KNN with distance = euclidean
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 4 of 5
Fit: KNN (k=15, distance=euclidean)
Predict: Using KNN with distance = euclidean
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
```

```
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Fold 5 of 5
 Fit: KNN (k=15, distance=euclidean)
 Predict: Using KNN with distance = euclidean
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Average CV MSE for k = 15 is 0.4887360075411709
 Testing k = 19
    Fold 1 of 5
 Fit: KNN (k=19, distance=euclidean)
 Predict: Using KNN with distance = euclidean
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Fold 2 of 5
 Fit: KNN (k=19, distance=euclidean)
 Predict: Using KNN with distance = euclidean
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
```

```
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=19, distance=euclidean)
Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=19, distance=euclidean)
Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 5 of 5
Fit: KNN (k=19, distance=euclidean)
Predict: Using KNN with distance = euclidean
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
```

```
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Average CV MSE for k = 19 is 0.49255148928453796
Finished Grid Search for euclidean
Best k (Euclidean): 7
   Fit: KNN (k=7, distance=euclidean)
   Predict: Using KNN with distance = euclidean
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
      Predicted 1300 samples
      Predicted 1400 samples
      Predicted 1500 samples
      Predicted 1600 samples
      Predicted 1700 samples
      Predicted 1800 samples
      Predicted 1900 samples


Distance: Manhattan
Grid Search for distance: manhattan
  Testing k = 1
     Fold 1 of 5
   Fit: KNN (k=1, distance=manhattan)
   Predict: Using KNN with distance = manhattan
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
```

```
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
 Fold 2 of 5
Fit: KNN (k=1, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
 Fold 3 of 5
Fit: KNN (k=1, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
 Fold 4 of 5
Fit: KNN (k=1, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
```

```
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
 Fold 5 of 5
Fit: KNN (k=1, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
   Average CV MSE for k = 1 is 0.60012115828744
Testing k = 3
 Fold 1 of 5
Fit: KNN (k=3, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
 Fold 2 of 5
Fit: KNN (k=3, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
```

```
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 3 of 5
Fit: KNN (k=3, distance=manhattan)
Predict: Using KNN with distance = manhattan
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 4 of 5
Fit: KNN (k=3, distance=manhattan)
Predict: Using KNN with distance = manhattan
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 5 of 5
Fit: KNN (k=3, distance=manhattan)
Predict: Using KNN with distance = manhattan
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
```

```
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Average CV MSE for k = 3 is 0.5086804823745356
Testing k = 5
   Fold 1 of 5
 Fit: KNN (k=5, distance=manhattan)
 Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Fold 2 of 5
 Fit: KNN (k=5, distance=manhattan)
 Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Fold 3 of 5
 Fit: KNN (k=5, distance=manhattan)
 Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
```

```
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 4 of 5
 Fit: KNN (k=5, distance=manhattan)
 Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 5 of 5
 Fit: KNN (k=5, distance=manhattan)
 Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Average CV MSE for k = 5 is 0.4882050392291152
Testing k = 7
  Fold 1 of 5
 Fit: KNN (k=7, distance=manhattan)
 Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
```

```
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 2 of 5
Fit: KNN (k=7, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=7, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=7, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
```

```
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Fold 5 of 5
 Fit: KNN (k=7, distance=manhattan)
 Predict: Using KNN with distance = manhattan
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Average CV MSE for k = 7 is 0.4863480978659928
 Testing k = 9
    Fold 1 of 5
 Fit: KNN (k=9, distance=manhattan)
 Predict: Using KNN with distance = manhattan
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Fold 2 of 5
 Fit: KNN (k=9, distance=manhattan)
 Predict: Using KNN with distance = manhattan
```

```
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=9, distance=manhattan)
Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=9, distance=manhattan)
Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 5 of 5
Fit: KNN (k=9, distance=manhattan)
Predict: Using KNN with distance = manhattan
```

```
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Average CV MSE for k = 9 is 0.4837643886862325
Testing k = 11
  Fold 1 of 5
 Fit: KNN (k=11, distance=manhattan)
 Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 2 of 5
 Fit: KNN (k=11, distance=manhattan)
 Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 3 of 5
```

```
Fit: KNN (k=11, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=11, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 5 of 5
Fit: KNN (k=11, distance=manhattan)
Predict: Using KNN with distance = manhattan
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Average CV MSE for k = 11 is 0.4809257289504811
```

```
Testing k = 15
   Fold 1 of 5
 Fit: KNN (k=15, distance=manhattan)
 Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Fold 2 of 5
 Fit: KNN (k=15, distance=manhattan)
 Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Fold 3 of 5
 Fit: KNN (k=15, distance=manhattan)
 Predict: Using KNN with distance = manhattan
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
```

```
      Predicted 1200 samples
    Fold 4 of 5
  Fit: KNN (k=15, distance=manhattan)
  Predict: Using KNN with distance = manhattan
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Fold 5 of 5
  Fit: KNN (k=15, distance=manhattan)
  Predict: Using KNN with distance = manhattan
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Average CV MSE for k = 15 is 0.4856087159121366
 Testing k = 19
    Fold 1 of 5
  Fit: KNN (k=19, distance=manhattan)
  Predict: Using KNN with distance = manhattan
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
```

```
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 2 of 5
Fit: KNN (k=19, distance=manhattan)
Predict: Using KNN with distance = manhattan
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 3 of 5
Fit: KNN (k=19, distance=manhattan)
Predict: Using KNN with distance = manhattan
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 4 of 5
Fit: KNN (k=19, distance=manhattan)
Predict: Using KNN with distance = manhattan
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
```

```
        Predicted 1000 samples
        Predicted 1100 samples
        Predicted 1200 samples
      Fold 5 of 5
    Fit: KNN (k=19, distance=manhattan)
    Predict: Using KNN with distance = manhattan
        Predicted 0 samples
        Predicted 100 samples
        Predicted 200 samples
        Predicted 300 samples
        Predicted 400 samples
        Predicted 500 samples
        Predicted 600 samples
        Predicted 700 samples
        Predicted 800 samples
        Predicted 900 samples
        Predicted 1000 samples
        Predicted 1100 samples
        Predicted 1200 samples
      Average CV MSE for k = 19 is 0.48717413211159366
Finished Grid Search for manhattan
Best k (Manhattan): 11
    Fit: KNN (k=11, distance=manhattan)
    Predict: Using KNN with distance = manhattan
        Predicted 0 samples
        Predicted 100 samples
        Predicted 200 samples
        Predicted 300 samples
        Predicted 400 samples
        Predicted 500 samples
        Predicted 600 samples
        Predicted 700 samples
        Predicted 800 samples
        Predicted 900 samples
        Predicted 1000 samples
        Predicted 1100 samples
        Predicted 1200 samples
        Predicted 1300 samples
        Predicted 1400 samples
        Predicted 1500 samples
        Predicted 1600 samples
        Predicted 1700 samples
        Predicted 1800 samples
        Predicted 1900 samples

Distance: Minkowski (p=3)
Grid Search for distance: minkowski
  Testing k = 1
```

```
  Fold 1 of 5
Fit: KNN (k=1, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 2 of 5
Fit: KNN (k=1, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=1, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
```

```
   Fold 4 of 5
 Fit: KNN (k=1, distance=minkowski)
 Predict: Using KNN with distance = minkowski
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Fold 5 of 5
 Fit: KNN (k=1, distance=minkowski)
 Predict: Using KNN with distance = minkowski
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
   Average CV MSE for k = 1 is 0.6316682655296975
 Testing k = 3
   Fold 1 of 5
 Fit: KNN (k=3, distance=minkowski)
 Predict: Using KNN with distance = minkowski
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
```

```
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 2 of 5
Fit: KNN (k=3, distance=minkowski)
Predict: Using KNN with distance = minkowski
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 3 of 5
Fit: KNN (k=3, distance=minkowski)
Predict: Using KNN with distance = minkowski
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 4 of 5
Fit: KNN (k=3, distance=minkowski)
Predict: Using KNN with distance = minkowski
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
```

```
   Predicted 1100 samples
   Predicted 1200 samples
 Fold 5 of 5
Fit: KNN (k=3, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
 Average CV MSE for k = 3 is 0.5154682173084904
Testing k = 5
 Fold 1 of 5
Fit: KNN (k=5, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
 Fold 2 of 5
Fit: KNN (k=5, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
```

```
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 3 of 5
Fit: KNN (k=5, distance=minkowski)
Predict: Using KNN with distance = minkowski
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 4 of 5
Fit: KNN (k=5, distance=minkowski)
Predict: Using KNN with distance = minkowski
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 5 of 5
Fit: KNN (k=5, distance=minkowski)
Predict: Using KNN with distance = minkowski
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
```

```
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Average CV MSE for k = 5 is 0.49037387319057296
Testing k = 7
  Fold 1 of 5
 Fit: KNN (k=7, distance=minkowski)
 Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 2 of 5
 Fit: KNN (k=7, distance=minkowski)
 Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 3 of 5
 Fit: KNN (k=7, distance=minkowski)
 Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
```

```
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Fold 4 of 5
 Fit: KNN (k=7, distance=minkowski)
 Predict: Using KNN with distance = minkowski
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Fold 5 of 5
 Fit: KNN (k=7, distance=minkowski)
 Predict: Using KNN with distance = minkowski
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
      Predicted 500 samples
      Predicted 600 samples
      Predicted 700 samples
      Predicted 800 samples
      Predicted 900 samples
      Predicted 1000 samples
      Predicted 1100 samples
      Predicted 1200 samples
    Average CV MSE for k = 7 is 0.4844860028733596
 Testing k = 9
    Fold 1 of 5
 Fit: KNN (k=9, distance=minkowski)
 Predict: Using KNN with distance = minkowski
      Predicted 0 samples
      Predicted 100 samples
      Predicted 200 samples
      Predicted 300 samples
      Predicted 400 samples
```

```
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 2 of 5
Fit: KNN (k=9, distance=minkowski)
Predict: Using KNN with distance = minkowski
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 3 of 5
Fit: KNN (k=9, distance=minkowski)
Predict: Using KNN with distance = minkowski
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
  Predicted 500 samples
  Predicted 600 samples
  Predicted 700 samples
  Predicted 800 samples
  Predicted 900 samples
  Predicted 1000 samples
  Predicted 1100 samples
  Predicted 1200 samples
 Fold 4 of 5
Fit: KNN (k=9, distance=minkowski)
Predict: Using KNN with distance = minkowski
  Predicted 0 samples
  Predicted 100 samples
  Predicted 200 samples
  Predicted 300 samples
  Predicted 400 samples
```

```
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 5 of 5
 Fit: KNN (k=9, distance=minkowski)
 Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
   Average CV MSE for k = 9 is 0.4875670170853498
Testing k = 11
  Fold 1 of 5
 Fit: KNN (k=11, distance=minkowski)
 Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 2 of 5
 Fit: KNN (k=11, distance=minkowski)
 Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
```

```
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=11, distance=minkowski)
Predict: Using KNN with distance = minkowski
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=11, distance=minkowski)
Predict: Using KNN with distance = minkowski
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 5 of 5
Fit: KNN (k=11, distance=minkowski)
Predict: Using KNN with distance = minkowski
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
```

```
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Average CV MSE for k = 11 is 0.49063564537851956
Testing k = 15
  Fold 1 of 5
 Fit: KNN (k=15, distance=minkowski)
 Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 2 of 5
 Fit: KNN (k=15, distance=minkowski)
 Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 3 of 5
 Fit: KNN (k=15, distance=minkowski)
 Predict: Using KNN with distance = minkowski
   Predicted 0 samples
```

```
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 4 of 5
 Fit: KNN (k=15, distance=minkowski)
 Predict: Using KNN with distance = minkowski
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Fold 5 of 5
 Fit: KNN (k=15, distance=minkowski)
 Predict: Using KNN with distance = minkowski
    Predicted 0 samples
    Predicted 100 samples
    Predicted 200 samples
    Predicted 300 samples
    Predicted 400 samples
    Predicted 500 samples
    Predicted 600 samples
    Predicted 700 samples
    Predicted 800 samples
    Predicted 900 samples
    Predicted 1000 samples
    Predicted 1100 samples
    Predicted 1200 samples
  Average CV MSE for k = 15 is 0.49593958869960586
Testing k = 19
  Fold 1 of 5
 Fit: KNN (k=19, distance=minkowski)
```

```
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 2 of 5
Fit: KNN (k=19, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 3 of 5
Fit: KNN (k=19, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 4 of 5
Fit: KNN (k=19, distance=minkowski)
```

```
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Fold 5 of 5
Fit: KNN (k=19, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
  Average CV MSE for k = 19 is 0.4993375809582264
Finished Grid Search for minkowski
Best k (Minkowski): 7
Fit: KNN (k=7, distance=minkowski)
Predict: Using KNN with distance = minkowski
   Predicted 0 samples
   Predicted 100 samples
   Predicted 200 samples
   Predicted 300 samples
   Predicted 400 samples
   Predicted 500 samples
   Predicted 600 samples
   Predicted 700 samples
   Predicted 800 samples
   Predicted 900 samples
   Predicted 1000 samples
   Predicted 1100 samples
   Predicted 1200 samples
```

```
Predicted 1300 samples
Predicted 1400 samples
Predicted 1500 samples
Predicted 1600 samples
Predicted 1700 samples
Predicted 1800 samples
Predicted 1900 samples

Running Linear Regression baseline
Baseline completed
Saving results to CSV file: stage1_results.csv
Done.
```

[ ]: