

Hyperparameter Optimization: Takeaways



by Dataquest Labs, Inc. - All rights reserved © 2021

Syntax

- Using Grid Search to find the optimal k value:

```
from sklearn.neighbors import KNeighborsRegressor
from sklearn.metrics import mean_squared_error
cols = ['accommodates', 'bedrooms', 'bathrooms', 'number_of_reviews']
hyper_params = [x for x in range(5)]
mse_values = list()
for value in hyper_params:
    knn = KNeighborsRegressor(n_neighbors=value, algorithm='brute')
    knn.fit(train_df[cols], train_df['price'])
    predictions = knn.predict(test_df[cols])
    mse = mean_squared_error(test_df['price'], predictions)
    mse_values.append(mse)
```

- Plotting to visualize the optimal k value:

```
features = ['accommodates', 'bedrooms', 'bathrooms', 'number_of_reviews']
hyper_params = [x for x in range(1, 21)]
mse_values = list()
for hp in hyper_params:
    knn = KNeighborsRegressor(n_neighbors=hp, algorithm='brute')
    knn.fit(train_df[features], train_df['price'])
    predictions = knn.predict(test_df[features])
    mse = mean_squared_error(test_df['price'], predictions)
    mse_values.append(mse)
plt.scatter(hyper_params, mse_values)
plt.show()
```

Concepts

- Hyperparameters are values that affect the behavior and performance of a model that are unrelated to the data. Hyperparameter optimization is the process of finding the optimal hyperparameter value.
- Grid search is a simple but common hyperparameter optimization technique, which involves evaluating the model performance at different k values and selecting the k value that resulted in the lowest error. Grid search involves:
 - Selecting a subset of the possible hyperparameter values.
 - Training a model using each of these hyperparameter values.
 - Evaluating each model's performance.
 - Selecting the hyperparameter value that resulted in the lowest error value.

- The general workflow for finding the best model is:
 - Selecting relevant features to use for predicting the target column.
 - Using grid search to find the optimal hyperparameter value for the selected features.
 - Evaluate the model's accuracy and repeat the process.

Resources

- [Difference Between Parameter and Hyperparameter](#)
- [Hyperparameter Optimization](#)

Takeaways by Dataquest Labs, Inc. - All rights reserved © 2021