**K-fold cross-validation** is a technique used to assess the performance of machine learning models. It helps to ensure that the model performs well on different subsets of data and isn't just overfitting to a particular part of the dataset.

**How it works:**

1. **Split the dataset into K subsets (or folds)**:
   * If you have a dataset with 1000 data points and you choose **K=5**, the data is split into 5 subsets, each containing 200 data points.
2. **Train and Test the Model K times**:
   * In each iteration, one of the K subsets is used as the **test set**, and the remaining K-1 subsets are combined to form the **training set**.
   * For example, in the first iteration, fold 1 will be used for testing, and the remaining folds (2, 3, 4, and 5) will be used to train the model.
   * In the second iteration, fold 2 is used for testing, and the remaining folds (1, 3, 4, and 5) are used for training.
   * This process repeats until each fold has been used as a test set exactly once.
3. **Evaluate the Model Performance**:
   * After K iterations, the model’s performance is averaged over all the test sets. For example, if you're measuring accuracy, you calculate the average accuracy from all K test sets.
   * This gives a more reliable estimate of the model’s performance, as it has been tested on different parts of the data.

**Why K-fold Cross-validation is useful:**

* **Reduces bias**: It uses the entire dataset for both training and testing. Each data point is used for testing exactly once, so the results are not biased toward any particular subset.
* **Helps with overfitting**: By using different parts of the data for training and testing, K-fold cross-validation helps to ensure that the model generalizes well to new data, not just the specific training data.
* **More reliable evaluation**: It gives a more robust estimate of how the model will perform when applied to real-world, unseen data.

**Example:**

Suppose you have a dataset with 1000 data points, and you want to use **5-fold cross-validation** (K=5). Here's how it would work:

1. Split the data into 5 parts (folds), each with 200 data points.
2. **1st fold**: Train the model on folds 2, 3, 4, and 5, and test it on fold 1.
3. **2nd fold**: Train the model on folds 1, 3, 4, and 5, and test it on fold 2.
4. **3rd fold**: Train the model on folds 1, 2, 4, and 5, and test it on fold 3.
5. **4th fold**: Train the model on folds 1, 2, 3, and 5, and test it on fold 4.
6. **5th fold**: Train the model on folds 1, 2, 3, and 4, and test it on fold 5.

After all 5 folds, you will have 5 performance metrics (e.g., accuracy), and you can calculate the average performance.

**In summary:**

K-fold cross-validation helps you get a better estimate of how well your model will perform on unseen data by testing it on multiple different subsets of the data and then averaging the results. This process helps to make sure the model is not just memorizing the training data but can generalize well.