The features are always randomly permuted at each split. Therefore, the best found split may vary, even with the same training data and max\_features=n\_features, if the improvement of the criterion is identical for several splits enumerated during the search of the best split. To obtain a deterministic behaviour during fitting, random\_state has to be fixed.

Các tính năng luôn được hoán vị ngẫu nhiên ở mỗi lần phân tách. Do đó, phần tách tốt nhất được tìm thấy có thể khác nhau, ngay cả với cùng một dữ liệu huấn luyện và max\_features=n\_features, nếu tiêu chí cải thiện giống hệt nhau đối với một số phần tách được liệt kê trong quá trình tìm kiếm phần tách tốt nhất. Để có được một hành vi xác định trong quá trình điều chỉnh, Random\_state phải được sửa.

**score**(*X*, *y*, *sample\_weight=None*)[[source]](https://github.com/scikit-learn/scikit-learn/blob/dc580a8ef/sklearn/base.py#L612)

Return the mean accuracy on the given test data and labels.

Trả về độ chính xác trung bình trên tập test và nhãn.

In multi-label classification, this is the subset accuracy which is a harsh metric since you require for each sample that each label set be correctly predicted.

Trong phân loại nhiều nhãn, đây là độ chính xác của tập hợp con, đây là một chỉ số khắc nghiệt vì bạn yêu cầu dự đoán chính xác từng tập hợp nhãn đối với từng mẫu.

**random\_state*int, RandomState instance or None, default=None***

Controls the random seed given to each Tree estimator at each boosting iteration. In addition, it controls the random permutation of the features at each split (see Notes for more details). It also controls the random splitting of the training data to obtain a validation set if n\_iter\_no\_change is not None. Pass an int for reproducible output across multiple function calls. See [Glossary](https://scikit-learn.org/stable/glossary.html#term-random_state).

Random\_stateint, RandomState instance hoặc Không, default=None Kiểm soát hạt giống ngẫu nhiên được cung cấp cho mỗi công cụ ước tính Cây ở mỗi lần lặp tăng cường. Ngoài ra, nó kiểm soát hoán vị ngẫu nhiên của các tính năng ở mỗi lần phân tách (xem Ghi chú để biết thêm chi tiết). Nó cũng kiểm soát việc phân tách ngẫu nhiên dữ liệu huấn luyện để có được bộ xác thực nếu n\_iter\_no\_change không phải là Không có. Truyền một int cho đầu ra có thể lặp lại qua nhiều lệnh gọi hàm.

* **Classification Accuracy:** It is one of the important parameters to determine the accuracy of the classification problems. It defines how often the model predicts the correct output. It can be calculated as the ratio of the number of correct predictions made by the classifier to all number of predictions made by the classifiers. The formula is given below:  
  Confusion Matrix in Machine Learning
* Dự đoán tần số chính xác của mô hình đầu ra.
* ***Misclassification rate*** *: It is also termed as Error rate, and it defines how often the model gives the wrong predictions. The value of error rate can be calculated as the number of incorrect predictions to all number of the predictions made by the classifier. The formula is given below:*Confusion Matrix in Machine Learning
* **Precision (sn):** It can be defined as the number of correct outputs provided by the model or out of all positive classes that have predicted correctly by the model, how many of them were actually true. It can be calculated using the below formula:  
  Confusion Matrix in Machine Learning
* **Recall:** It is defined as the out of total positive classes, how our model predicted correctly. The recall must be as high as possible.  
  Confusion Matrix in Machine Learning
* **F-measure:** If two models have low precision and high recall or vice versa, it is difficult to compare these models. So, for this purpose, we can use F-score. This score helps us to evaluate the recall and precision at the same time. The F-score is maximum if the recall is equal to the precision. It can be calculated using the below formula:  
  Confusion Matrix in Machine Learning

Other important terms used in Confusion Matrix:

* **Null Error rate:** It defines how often our model would be incorrect if it always predicted the majority class. As per the accuracy paradox, it is said that "*the best classifier has a higher error rate than the null error rate.*"
* **ROC Curve:** The ROC is a graph displaying a classifier's performance for all possible thresholds. The graph is plotted between the true positive rate (on the Y-axis) and the false Positive rate (on the x-axis).

ROC là biểu đồ hiển thị hiệu suất của bộ phân loại đối với tất cả các ngưỡng có thể. Biểu đồ được vẽ giữa tỷ lệ dương thực (trên trục Y) và tỷ lệ dương sai (trên trục x).