GridSearchCV and RandomizedSearchCV

ML Assignment - 9 by Athika Fatima - 101502209

GridSearchCV

GridSearchCV is a hyperparameter tuning technique in machine learning that systematically searches through a specified parameter grid. It evaluates all possible combinations of parameters to determine the best configuration based on cross-validation performance.

Key Features:

- Exhaustive search over a predefined parameter space.
- Uses cross-validation to assess each combination.
- Guarantees finding the best combination (within the grid).
- Can be computationally expensive for large parameter spaces.

RandomizedSearchCV

RandomizedSearchCV is an alternative hyperparameter tuning method that samples a fixed number of parameter settings from a given distribution, rather than exhaustively searching all combinations.

Key Features:

- Selects a random subset of hyperparameters rather than evaluating all combinations.
- More computationally efficient, especially for large parameter spaces.
- Can still yield excellent results if appropriately configured.
- Allows defining distributions for hyperparameters rather than fixed values.

Why is RandomizedSearchCV Needed When GridSearchCV Already Exists?

Although GridSearchCV guarantees finding the best combination, it becomes impractical for models with multiple hyperparameters, especially when:

- The number of combinations is too large.
- Computing resources are limited.
- Some hyperparameters contribute more significantly to performance than others, making exhaustive search unnecessary.

RandomizedSearchCV helps in such cases by efficiently exploring a broad parameter space without requiring an exhaustive search.

When to Use Which Cross-Validation Approach?

Use GridSearchCV When:

- The parameter space is small.
- You need precise tuning and have sufficient computational resources.
- The dataset is relatively small, making exhaustive evaluation feasible.

Use RandomizedSearchCV When:

- The parameter space is large and complex.
- Computational resources are limited.
- Approximate tuning is sufficient for achieving good performance.
- You want to quickly identify promising hyperparameter values before refining them further.

Can GridSearchCV and RandomizedSearchCV Be Used Together?

Yes, both methods can be used in combination to balance efficiency and precision. A common approach is:

1. **Use RandomizedSearchCV First:** This helps in identifying a promising region of the hyperparameter space efficiently.
2. **Follow Up with GridSearchCV:** Once a good range is identified, use GridSearchCV to fine-tune the best-performing hyperparameters within that range.

This two-step approach optimizes both exploration and precision, making it an effective strategy for hyperparameter tuning.