1. Filter Methods
2. Wrapper Methods
3. Embedded Methods

**Filter Methods:**

1. **Information Gain:**
   * Measures the reduction in entropy when a feature is included.
   * Used for feature selection by evaluating information gain of each variable in the context of the target variable.
2. **Chi-square Test:**
   * Used for categorical features.
   * Calculates Chi-square between each feature and the target.
   * Selects features with the best Chi-square scores.
3. **Fisher’s Score:**
   * A supervised feature selection method based on Fisher's score.
   * Ranks variables based on Fisher’s score in descending order.
4. **Correlation Coefficient:**
   * Uses correlation (e.g., Pearson Correlation) to measure linear relationships.
   * Good variables correlate highly with the target.
   * Handles multicollinearity by dropping variables with lower correlation coefficients.
5. **Variance Threshold:**
   * Removes features with low variance.
   * Assumes features with higher variance contain more useful information.
6. **Mean Absolute Difference (MAD):**
   * Computes absolute difference from the mean value.
   * Scaled variant; higher MAD indicates higher discriminatory power.
7. **Dispersion Ratio:**
   * Uses arithmetic mean (AM) and geometric mean (GM) to calculate dispersion ratio.
   * Higher dispersion implies a more relevant feature.

**Wrapper Methods:**

1. **Forward Feature Selection:**
   * Iterative method starting with one feature and adding the best-performing features.
   * Continues until a preset criterion is achieved.
2. **Backward Feature Elimination:**
   * Starts with all features and removes the least important one in each iteration.
   * Continues until a preset criterion is achieved.
   * Also known as Sequential Feature Selection.
3. **Exhaustive Feature Selection:**
   * Brute-force evaluation of each feature subset.
   * Tries every possible combination of variables and returns the best-performing subset.
4. **Recursive Feature Elimination (RFE):**
   * Uses an external estimator that assigns weights to features.
   * Selects features by recursively considering smaller sets until the desired number is reached.

**Embedded Methods:**

1. **LASSO Regularization (L1):**
   * Adds a penalty to model parameters to reduce overfitting.
   * Can shrink some coefficients to zero, effectively removing features.
2. **Random Forest Importance:**
   * Aggregates decision trees and ranks features by their contribution to impurity reduction.
   * Pruning trees below a certain node creates a subset of the most important features.