Feature Selection:

- Feature selection techniques in machine learning involve selecting the most relevant features or variables from a dataset, which helps to reduce the dimensionality of the data and improve model performance.

Need of Feature Selection:

- Feature selection reduces the dimensionality of the data, and it reducing the risk of overfitting.
- It removes irrelevant or redundant features that can negatively impact model performance and accuracy.
- It help to reduce training time and computational costs.
- It can generalize well to new data.

Types of Feature Selections:

> 1. Forward Feature Selection and Backward Elimination
> 2. Filter methods
> 3. Wrapper methods
> 4. Embedded methods
> 5. Principal Component Analysis (PCA)
> 6. Recursive Feature Elimination (RFE)
> 7. Lasso Regression
> 8. Genetic Algorithms
> 9. Univariate Feature Selection

1. Forward Feature Selection and Backward Elimination:

- **Forward Feature Selection** : Forward Feature Selection is a feature selection technique that iteratively builds a model by adding one feature at a time.
- **Backward Elimination** : Backward Elimination systematically removes features from the model one at a time, evaluating the impact on model performance until no further improvement is observed.

1.1. Step-by-Step Process of Forward Feature Selection:

Example Dataset:

x1	x2	x3	y
2	3	1	10
4	2	3	20
5	6	2	30
7	8	4	40

x1	x2	x3	y
1	3	5	15

Goal:

- Use Forward Feature Selection to determine which features (x1, x2, x3) best predict the target variable y.

1. Start with an Empty Model

- Initially, the model has no features.
- Set of selected features S=∅.

2. Evaluate Each Feature Individually

- Train a model using each feature separately and compute its performance.

Feature	Model Performance (R^2)
x1	0.65
x2	0.72
x3	0.50

- The best-performing feature is x2 (R^2 = 0.72), so we add it to the selected set. Updated set: S={x2}.

3. Add the Next Best Feature

- Train models with S + each remaining feature:

Feature Set (S + new feature)	Model Performance (R^2)
{x2, x1}	0.80
{x2, x3}	0.75

- The best-performing combination is {x2, x1} (R^2 = 0.80), so we add x1.
- Updated set: S={x2,x1}.

4. Evaluate the Last Feature

- Train a model with all three features:

Feature Set (S + x3)	Model Performance (R^2)
{x2, x1, x3}	0.78

- Since adding x3 decreases performance (R^2 drops from 0.80 to 0.78), we do not add it.

5. Final Selected Features

- The optimal feature set is {x2, x1}.