Feature selection methods

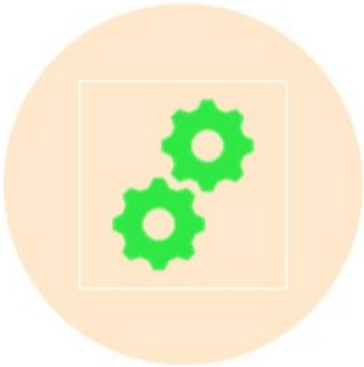
Filter

Wrapper methods

Embedded methods

Pre-reqs

PYTHON



MACHINE LEARNING MODELS

EVALUATION AND METRICS IN ML

Describe wrapper method

type of feature selection technique in machine learning that involves selecting subsets of features and evaluating their performance using a specific model.

wrapper methods use the performance of a chosen model as a criterion for feature selection.

Wrapper methods are computationally more expensive compared to filter methods, but they can lead to better models

Example

01

Dataset: Breast Cancer Dataset

02

Suppose we have following features:

- Mean radius
- Mean texture
- Mean smoothness
- ... (other features)

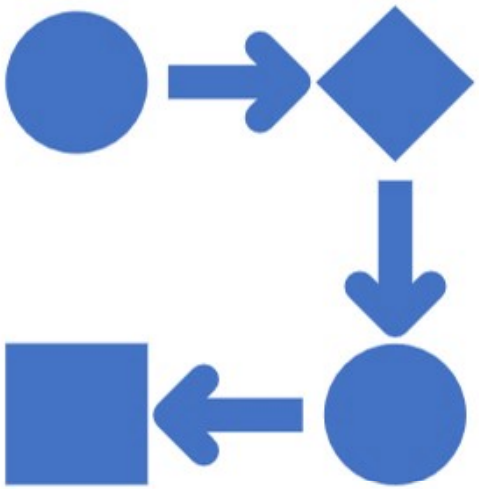
- Total of n=30 features

03

Objective : to find best k number of features

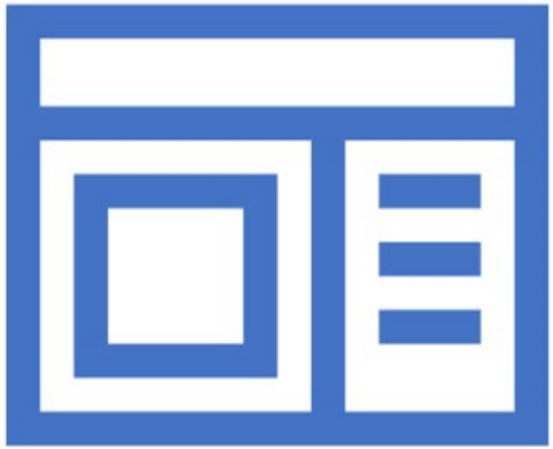
steps

- **1. Initialization**:
  - Start with the full set of features: [Mean radius, Mean texture, Mean smoothness, ...].
- **2. Model Training and Evaluation**:
  - Train a model (e.g., a support vector machine) using the current feature subset.
  - Evaluate the model's performance through cross-validation, obtaining an accuracy score of 0.92.
- **3. Feature Subset Update**:
  - Identify the least important feature, let's say it's "Mean smoothness," and remove it from the subset.
  - Updated feature subset: [Mean radius, Mean texture, ...].



… next

- 4. **Iteration**
  - Repeat steps 2-3 until the desired number of features is reached. For this example, let's say we want to keep 3 features.
- 5. **Final Model**
  - Train the final model using the selected feature subset [Mean radius, Mean texture, ...] on the entire dataset (rows).



Type of wrapper method - Forward Selection

- **Description:**
  - Start with an empty set of features and iteratively add features that improve model performance until a certain criterion is met.
- **Process:**
  - Begin with an empty set.
  - Iteratively add the most important feature not in the current subset until a stopping criterion is met.

example of forward selection using PIMA dataset

- Assuming our initial dataset has the following columns:
- **Glucose, BMI, Age, Pregnancies, BloodPressure, SkinThickness, Insulin, DiabetesPedigreeFunction,**
- and the target variable **Outcome.**

Iteration 1:

- Train individual models using features
  - [Pregnancies],
  - [Glucose],
  - [BloodPressure],
  - [SkinThickness],
  - [Insulin],
  - [BMI],
  - [DiabetesPedigreeFunction],
  - and [Age].
- Evaluate model performance for each feature.

  Suppose [Glucose] provides the best performance.

Iteration 1 (Update):

Feature subset: [**Glucose**].

Iteration 1 – (performance)

Run	Selected Feature	Model Performance (e.g., Accuracy)
1	Glucose	**0.75**
2	BMI	0.70
3	Age	0.72
4	Pregnancies	0.65
5	BloodPressure	0.68
6	SkinThickness	0.62
7	Insulin	0.70
8	DiabetesPedigreeFunction	0.71

Iteration 2

- Train models using features
  - [Glucose, Pregnancies],
  - [Glucose, BloodPressure],
  - [Glucose, SkinThickness],
  - [Glucose, Insulin],
  - [Glucose, BMI],
  - [Glucose, DiabetesPedigreeFunction], and
  - [Glucose, Age].
- Evaluate model performance for each feature pair.
- Suppose [Glucose , BMI] provides the best improvement.

Iteration 2 (Update):

Feature subset: [Glucose, BMI].

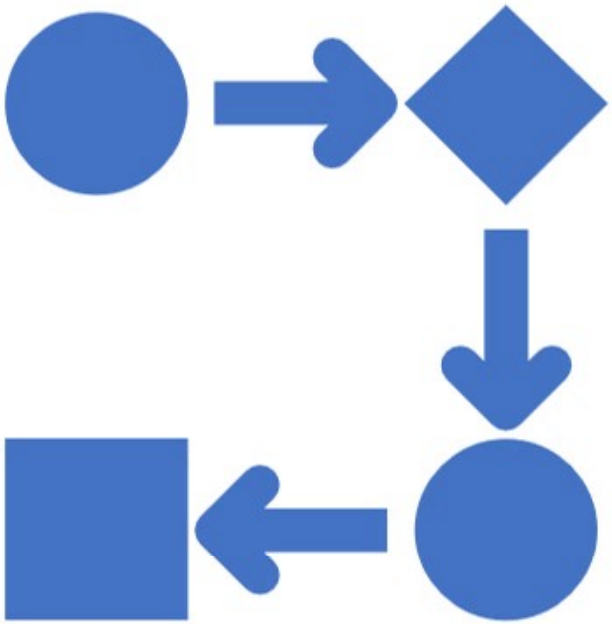
… next

Repeat:

Continue the process, adding features one by one based on their impact on model performance.

Stopping Criterion:

Stop when reaching a predefined number of features.



Wrapper method - Backward Elimination

Start with the full set of features

iteratively remove the least important features until a certain criterion is met.

Example

- Let's consider an example with a dataset containing the features X1, X2, X3, and X4.
- The goal is to predict a binary outcome Y.

Steps for Backward Elimination

1. Initialization:	Start with the features X1, X2, X3, X4.
2. Model Training and Evaluation (Iteration 1):	Train a model using all features (X1, X2, X3, X4). Evaluate the model's performance.

… next

3. Feature Removal (Iteration 1):

- Identify the least important feature (e.g., based on p-values from statistical tests).
- Suppose X3 is identified as the least important feature.
- Remove X3 from the feature set.

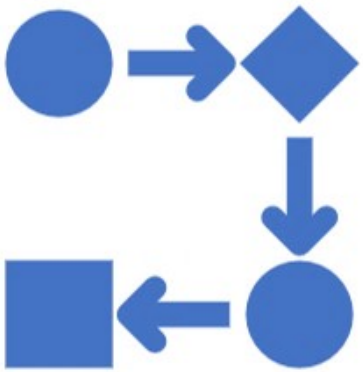
4. Model Training and Evaluation (Iteration 2):

- Train a new model using the updated set of features (X1, X2, X4).
- Evaluate the model's performance.

Step backward feature selection

Step	Descp	Col A	Col B	Col C	Col D	ML stats
1	Start with all columns	A	B	C	D	79%
2	Run with all the subsets	A	B	C		75
			B	C	D	72
		A		C	D	81
		A	B		D	67
3		A		C		76
				C	D	83
		A			D	77

… next



- **5. Repeat:**
  - Continue the process by identifying and removing the least important feature in each iteration.
  - Iteration 2: Identify and remove the least important feature.
  - Iteration 3: Identify and remove the least important feature.

… next

6. Stopping Criterion:

1. Stop when reaching a predefined number of features or achieving satisfactory model performance.

7. Final Model:

1. Use the selected feature subset to train the final model on the entire dataset.

Wrapper method

Recursive Feature Elimination (RFE)

Workout example

- For simplicity, we'll use a hypothetical dataset
- with four features (X1, X2, X3, X4) and a binary target variable (Y).

steps

Step 1: Initialization

- Start with the full set of features: X1, X2, X3, X4.

Step 2: Model Training and Evaluation (Iteration 1)

- Train a model using all features (X1, X2, X3, X4) and evaluate its performance.
- Suppose the model performance metrics are as follows:
  - Accuracy: 0.75
  - Feature Importance Ranking: X1 > X2 > X3 > X4

… next

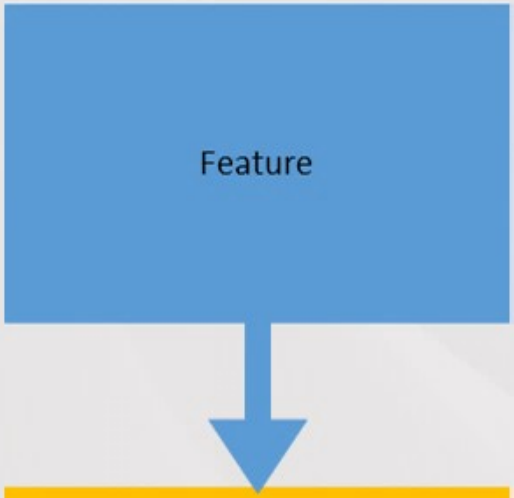
Step 3: Feature Elimination (Iteration 1)

- Identify the least important feature based on importance ranking or other criteria. In this case, suppose X4 is the least important.
- Remove X4 from the feature set: X1, X2, X3.

Step 4: Model Training and Evaluation (Iteration 2)

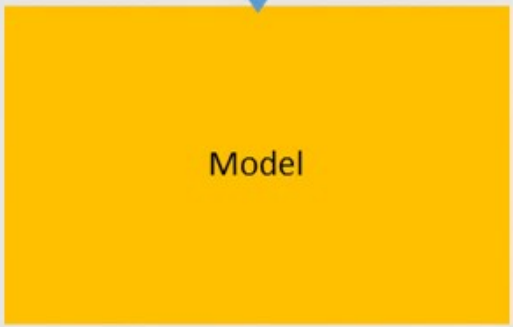
- Train a new model using the updated set of features (X1, X2, X3) and evaluate its performance.
- Suppose the model performance metrics are as follows:
- Accuracy: 0.78
- Feature Importance Ranking: X1 > X2 > X3

… next



Elimination (Iteration 2)

- Identify the least important feature. Suppose X2 is now identified as the least important.
- Remove X2 from the feature set: X1, X3.



Training and Evaluation (Iteration 3)

- Train a new model using the updated set of features (X1, X3) and evaluate its performance.
- Suppose the model performance metrics are as follows:
- Accuracy: 0.82
- Feature Importance Ranking: X1 > X3

Recursive Feature elimination (RFE)

Step	Descp	Col A	Col B	Col C	Col D	ML stats
1	Start with all columns	A	B	C	D	Calculate the score {coeff or feature importance} for each feature Remove the feature with lowest score
2	Run with all the subsets		B	C	D	Calculate the score {coeff or feature importance} for each feature Remove the feature with lowest score
3	Final Selection		B	C		

Thanks !!

- Next :
- Feature selection -> Embedded methods

(very important methods for ML/DL)