After,
data cleaning,
data preparation

Dimensionality reduction

(new Combination of attributes)

- Convert high dimension to low dimension feature space.
- To simplify the classification and regression for better for predictive modelling
- reducing no. of input feature in training sample

eg:

car mileage	car age	Target

=>

car mileage & age	target

- PCA (principle component analysis)
- Singular value decomposition

Different from each other but Common in to reduce the no. of attributes in dataset.

feature selection

(Scoring to each feature)

include & exclude data in data without changing

- using scoring and statistic method to select feature which (to keep) and (to delete) feature

eg:

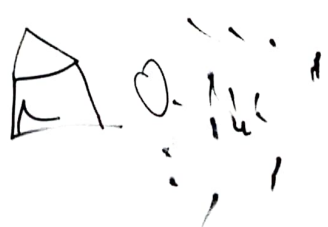
car name	car mileage	car age	car

include and

car mileage	car age	car type	

Algorithm

- Filter method
- Wrapper method
- Embedded method



fter,
a cleaning,
ta preparation

Dimensionality reduction

(new Combination of attributes)

- Convert high dimension to low dimension feature space.
- To simplify the classification and regression for better for predictive modelling
- reducing no. of input feature in training sample

eg:

car mileage	car age	Target

=>

car mileage & age	target

- PCA (principle component analysis)
- Singular value decomposition

Different from each other but Common in to reduce the no. of attributes in dataset.

(Scoring to each feature)

feature selection

- include & exclude data in dataset without changing them
- using scoring and statistic method to select feature which (to keep) and (to delete) feature

eg:

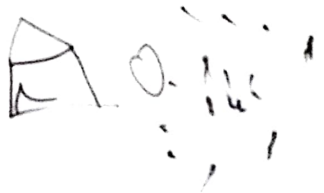
car name	car mileage	car age	car type	target

include and exclude

car mileage	car age	car type	target

Algorithm

- Filter method
- Wrapper method
- Embedded method



- Convert high dimension to low dimension feature space.
- To simplify the classification and regression for better for predictive modelling
- reducing no. of input feature in training sample

eg:

car milege	car age	Target

=>

car milege & age	target

Singular value decomposition

nciple
alysis)

Feature selection (Scoring to each feature)

(include & exclude data in dataset without changing them)

- using scoring and statistic method to select feature which (to keep) and (to delete) feature

eg:

car name	car milege	car age	car type	target

include and exclude

car milege	car age	car type	target

Algorithm

Filter method

- Apply a statistical measure to assign a scoring to each feature
- method are univariate

Some filter method

- Chi Squared test
- Information gain
- Correlation Coefficient Score

Wrapper method (recursive feature elimination algo)

- Selection of set of features as a search problem
- where,
  different combination are prepared, evaluated and compared to other combination of feature.
- assign a score based on model accuracy.

- Some wrapper method
  - Random hill climbing algorithm
  - heuristic like forward and backward pass.

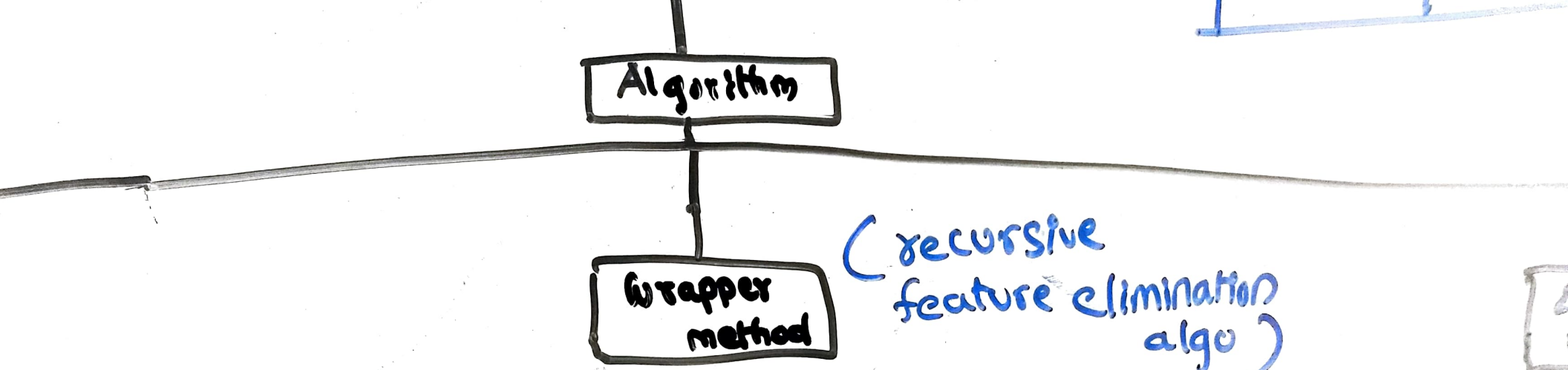
Embedded method

Filter method

- Apply a statistical measure to assign a scoring to each feature
- method are univariate

Some filter method

- Chi Squared test
- Information gain
- Correlation Coefficient Score



(recursive feature elimination algo)

assign a

- Selection of set of features as a search problem
  - where,
    different combination are prepared, evaluated and compared to other combination of feature.
  - assign a score based on model accuracy.
- Some wrapper method
  - Random hill climbing algorithm
  - heuristic like forward and backward pass.

eg:

Car mileage	car age	Target

=>

Car mileage & age	target

Singular value decomposition

Feature selection

- using scoring and statistic method to select feature which (to keep) and (to delete) feature

eg:

Car name	Car mileage	car age	Car type	target

↓ include and exclude

Car mileage	car age	car type	target

Algorithm

r method

Apply a statistical measure to assign a Scoring to each feature

method are univariate

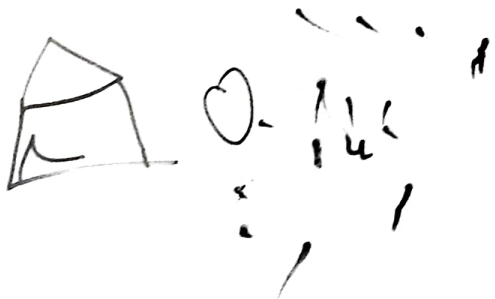
r method

squared test

rmation gain

lation Coefficient

e



Wrapper method

- Selection of set of features as a search problem
- where,
  different combination are prepared, evaluated and compared to other combine feature.
- assign a score based on mod accuracy.

- **Some wrapper method**
  - Random hill climbing alg
  - heuristic like forward pass.

Embedded method

- learn which feature is best contribute to accuracy of model while model being created

- **Common method:**
  - Regularization method (penalization method)
    - model toward low complexity
- **Example:**
  regularization algo are
  - LASSO
  - Elastic Net
  - Ridge Regression

Embedded method

- learn which feature is best contribute to accuracy of model while model being created

- Common method :

- Regularization method (penalization method)
  - model toward low complexity

- Example :

  regularization algo are
  - LASSO
  - Elastic Net
  - Ridge Regression

After,
data cleaning,
data preparation

Dimensionality reduction (new Combination of attributes)

- Convert high dimension to low dimension feature space,
- To simplify the classification and regression for better for predictive modelling
- reducing no. of input feature in training sample

eg:

Car milege	car age	Target

=>

car milage & age	target

- PCA (principle component analysis)
- Singular value decomposition

Different from each other but Common in to reduce the no. of attributes in dataset.

Feature selection (Scoring to each feature)

(include & exclude data in dataset without changing them)

- using scoring and statistic method to select feature which (to keep) and (to delete) feature

eg.

Car name	carmilage	car age	cartype	target

↓ include and exclude

car milage	car age	cartype	target

Algorithm

Filter method

- Apply a statistical measure to assign a scoring to each feature
- method are univariate

Some filter method

- Chi Squared test
- Information gain
- Correlation Coefficient Score

Wrapper method

- Selection of set of features as a search problem
- where,
  different combination are prepared, evaluated and compared to other combine feature.
- assign a score based on mod accuracy.

- Some wrapper method
  - Random hill climbing alg
  - heuristic like forward pass.

Embedded method

- learn which feature is best contribute to accuracy of model while model being created

- Common method:
  - Regularization method (penalization method)
    - model toward low complexity
- Example:
  regularization algo are
  - LASSO
  - Elastic Net
  - Ridge Regression

Feature Selection

Feature selection is a process where you automatically select those features in your data that contribute most to the prediction variable or output in which you are interested.

Having irrelevant features in your data can decrease the accuracy of many models, especially linear algorithms like linear and logistic regression.

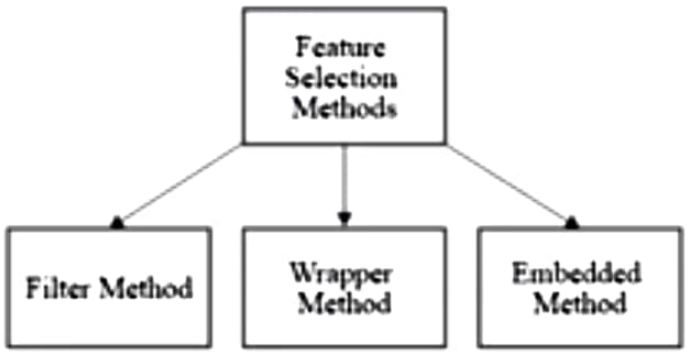
Three benefits of performing feature selection before modeling your data are:

- **Reduces Overfitting**: Less redundant data means less opportunity to make decisions based on noise.
- **Improves Accuracy**: Less misleading data means modeling accuracy improves.
- **Reduces Training Time**: Less data means that algorithms train faster.

You can learn more about feature selection with scikit-learn in the article Feature selection.

Feature Selection Methods

- Filter Methods
- Wrapper Methods
- Embedded Methods



Set of all Features → **Selecting the Best Subset** → **Learning Algorithm** → **Performance**

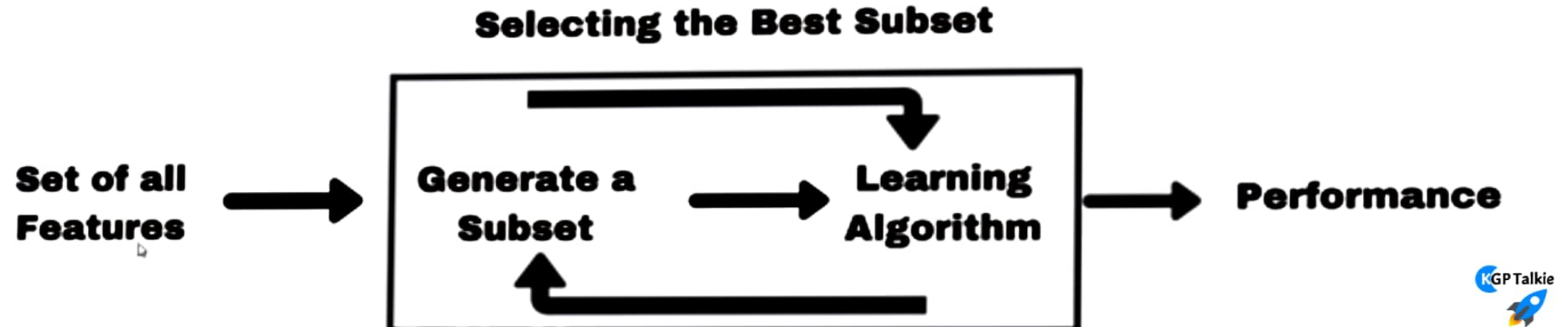
Filter Methods

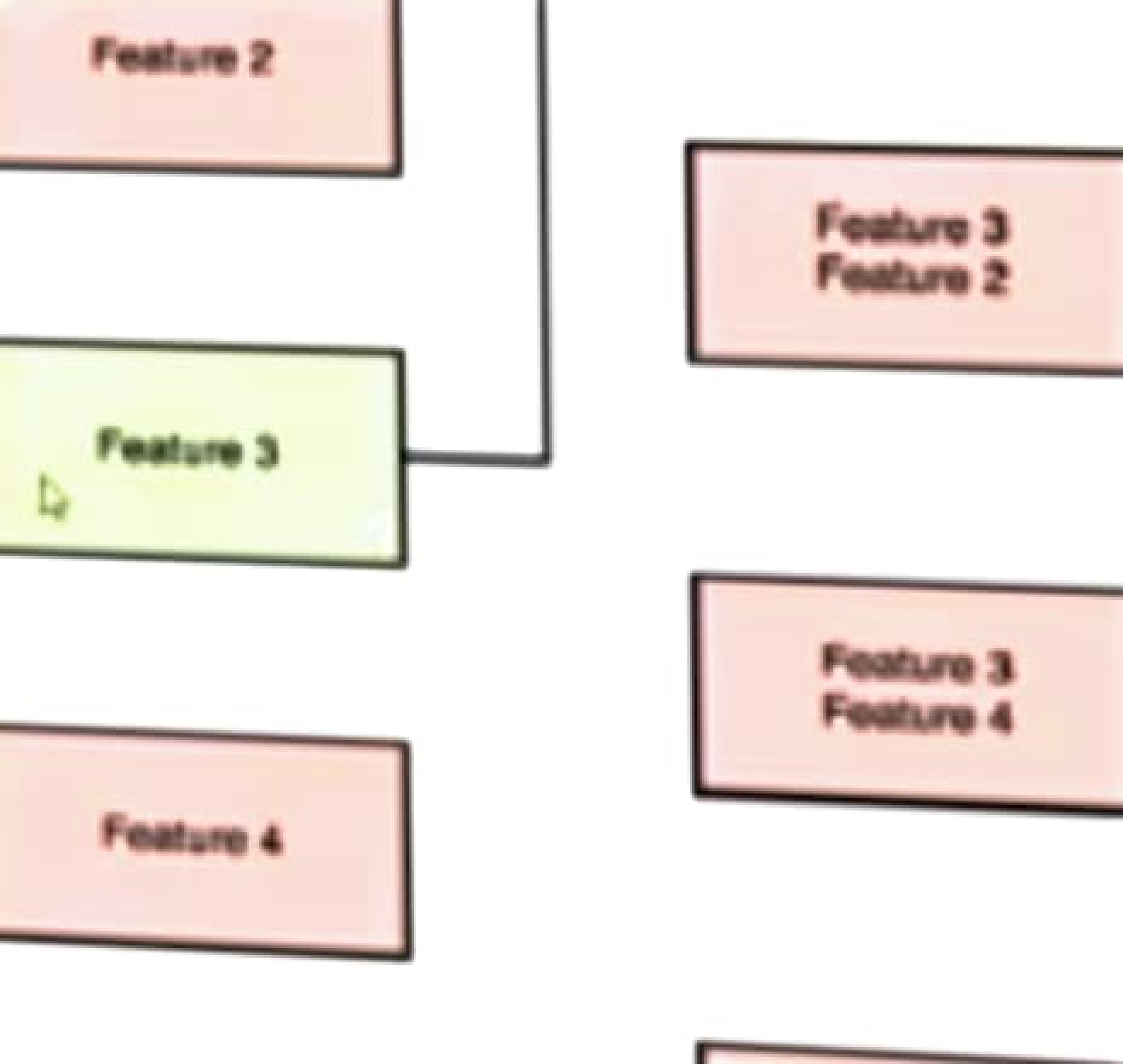
- Single Factor Analysis
- Uses Statistical algorithms like Chi-Square | Fisher Score | Anova | Mutual Information | Variance
- Uses Individual Feature Predictive Power
- Less computationally expensive
- Uses correlation power on feature to target space
- Useful in quick screening

Wrapper Methods

- Use combinations of variables to determine predictive power
- Find the best combination of variables
- Computationally expensive than filter method
- Perform better than filter method
- Not recommended on high number of features

- Subset Selection (Exhaustive Feature Selection)
- Forward Step Selection
- Backward Step Selection (Recursive Feature Selection)




Feature 2
Feature 3
Feature 3
Feature 2
Feature 3
Feature 4
Feature 4

Backward Step Selection (Recursive

Subset Selection (Exhaustive Feature Selection)

- fits the model with each possible combinations of N features.
- requires massive computational power

(Y = B0 , Y = B0 + B1*X1 , Y = C0 + C1*X2 , Y = D0 + D1*X1 + D2*X2)

- Use test error to evaluate model perfo

Embedded Methods

- Faster than wrapper methods
- More accurate than filter method
- Perform feature selection during algorithm implementation thus embedded in it

- LASSO Regression
- RIDGE Regression
- Tree Importance