Feature Selection : Backward Feature Elimination

Common Dimensionality Reduction Techniques

- Missing value ratio
- Low variance
- High correlation
- Backward feature elimination
- Forward feature selection

Common Dimensionality Reduction Techniques

- Missing value ratio
- Low variance
- High correlation
- Backward feature elimination
- Forward feature selection

Common Dimensionality Reduction Techniques

- Missing value ratio
- Low variance
- High correlation
- Backward feature elimination
- Forward feature selection

Feature Selection : Backward Feature Elimination

ID	Calories_burnt	Gender	Plays_Sport?	Fitness_Level
1	121	M	Yes	Fit
2	230	M	No	Fit
3	342	F	No	Unfit
4	70	M	Yes	Fit
5	278	F	Yes	Unfit
6	146	M	Yes	Fit
7	168	F	No	Unfit
8	231	F	Yes	Fit
9	150	M	No	Fit
10	190	F	No	Fit

Feature Selection : Backward Feature Elimination

- No missing values in the dataset
- Variance of the variables is high
- Low correlation between the independent variables

Backward Feature Elimination

Steps to perform Backward Feature Elimination

1. Train the model using all the variables (n)

Steps to perform Backward Feature Elimination

1. Train the model using all the variables (n)

ID	Calories_burnt	Gender	Plays_Sport?	Fitness_Level
1	121	M	Yes	Fit
2	230	M	No	Fit
3	342	F	No	Unfit
4	70	M	Yes	Fit
5	278	F	Yes	Unfit
6	146	M	Yes	Fit
7	168	F	No	Unfit
8	231	F	Yes	Fit
9	150	M	No	Fit
10	190	F	No	Fit

Steps to perform Backward Feature Elimination

1. Train the model using all the variables (n)

2. Calculate the performance of the model

Steps to perform Backward Feature Elimination

1. Train the model using all the variables (n)

2. Calculate the performance of the model

Accuracy = 92%

Steps to perform Backward Feature Elimination

1. Train the model using all the variables (n)

2. Calculate the performance of the model
3. Eliminate a variable, train the model on remaining variables (n-1)

Steps to perform Backward Feature Elimination

ID	Calories_burnt	Gender	Plays_Sport?	Fitness_Level
1	121	M	Yes	Fit
2	230	M	No	Fit
3	342	F	No	Unfit
4	70	M	Yes	Fit
5	278	F	Yes	Unfit
6	146	M	Yes	Fit
7	168	F	No	Unfit
8	231	F	Yes	Fit
9	150	M	No	Fit
10	190	F	No	Fit

Steps to perform Backward Feature Elimination

1. Train the model using all the variables (n)
2. Calculate the performance of the model
3. Eliminate a variable, train the model on remaining variables (n-1)
4. Calculate the performance of the model on new data

Steps to perform Backward Feature Elimination

ID	Calories_burnt	Gender	Plays_Sport?	Fitness_Level
1	121	M	Yes	Fit
2	230	M	No	Fit
3	342	F	No	Unfit
4	70	M	Yes	Fit
5	278	F	Yes	Unfit
6	146	M	Yes	Fit
7	168	F	No	Unfit
8	231	F	Yes	Fit
9	150	M	No	Fit
10	190	F	No	Fit

Accuracy =
90%

Steps to perform Backward Feature Elimination

ID	Calories_burnt	Gender	Plays_Sport?	Fitness_Level
1	121	M	Yes	Fit
2	230	M	No	Fit
3	342	F	No	Unfit
4	70	M	Yes	Fit
5	278	F	Yes	Unfit
6	146	M	Yes	Fit
7	168	F	No	Unfit
8	231	F	Yes	Fit
9	150	M	No	Fit
10	190	F	No	Fit

Steps to perform Backward Feature Elimination

ID	Calories_burnt	Gender	Plays_Sport?	Fitness_Level
1	121	M	Yes	Fit
2	230	M	No	Fit
3	342	F	No	Unfit
4	70	M	Yes	Fit
5	278	F	Yes	Unfit
6	146	M	Yes	Fit
7	168	F	No	Unfit
8	231	F	Yes	Fit
9	150	M	No	Fit
10	190	F	No	Fit

Accuracy =
91.6%

Steps to perform Backward Feature Elimination

ID	Calories_burnt	Gender	Plays_Sport?	Fitness_Level
1	121	M	Yes	Fit
2	230	M	No	Fit
3	342	F	No	Unfit
4	70	M	Yes	Fit
5	278	F	Yes	Unfit
6	146	M	Yes	Fit
7	168	F	No	Unfit
8	231	F	Yes	Fit
9	150	M	No	Fit
10	190	F	No	Fit

Steps to perform Backward Feature Elimination

ID	Calories_burnt	Gender	Plays_Sport?	Fitness_Level
1	121	M	Yes	Fit
2	230	M	No	Fit
3	342	F	No	Unfit
4	70	M	Yes	Fit
5	278	F	Yes	Unfit
6	146	M	Yes	Fit
7	168	F	No	Unfit
8	231	F	Yes	Fit
9	150	M	No	Fit
10	190	F	No	Fit

Accuracy =
88%

Steps to perform Backward Feature Elimination

1. Train the model using all the variables (n)

2. Calculate the performance of the model

3. Eliminate a variable, train the model on remaining variables

   (n-1)

4. Calculate the performance of the model on new data
5. Identify the eliminated variable which does not impact the performance much

Steps to perform Backward Feature Elimination

Accuracy using all the variables = 92%

Steps to perform Backward Feature Elimination

Accuracy using all the variables = 92%

Variable_dropped	Accuracy
Calories_burnt	90%
Gender	91.60%
Plays_Sport?	88%

Steps to perform Backward Feature Elimination

Accuracy using all the variables = 92%

Variable_dropped	Accuracy
Calories_burnt	90%
Gender	91.60%
Plays_Sport?	88%

Steps to perform Backward Feature Elimination

Accuracy using all the variables = 92%

Variable_dropped
Gender

Variable_dropped	Accuracy
Calories_burnt	90%
Gender	91.60%
Plays_Sport?	88%

Steps to perform Backward Feature Elimination

1. Train the model using all the variables (n)

2. Calculate the performance of the model

3. Eliminate a variable, train the model on remaining variables

   (n-1)

4. Calculate the performance of the model on new data
5. Identify the eliminated variable which does not impact the performance much

6. Repeat until no more variables can be dropped

Analytics Vidhya
Learn everything about analytics

Thank
You!

Analytics Vidhya
Learn everything about analytics