Feature Selection: 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

Common Dimensionality Reduction Techniques

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

Feature Selection: Forward Feature Selection

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

Forward Feature Selection

Steps to perform Forward Feature Selection

1. Train n model using each feature (n) individually and check the performance

Steps to perform Forward Feature Selection

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 Forward Feature Selection

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 =
87%

Steps to perform Forward Feature Selection

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 Forward Feature Selection

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 =
80%

Steps to perform Forward Feature Selection

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 Forward Feature Selection

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 = 85%

Steps to perform Forward Feature Selection

1. Train n model using each feature (n) individually and check the performance
2. Choose the variable which gives the best performance

Steps to perform Forward Feature Selection

1. Train n model using each feature (n) individually and check the performance
2. Choose the variable which gives the best performance

Variable used	Accuracy
Calories_burnt	87.00%
Gender	80.00%
Plays_Sport?	85.00%

Steps to perform Forward Feature Selection

1. Train n model using each feature (n) individually and check the performance
2. Choose the variable which gives the best performance

Variable used	Accuracy
Calories_burnt	87.00%
Gender	80.00%
Plays_Sport?	85.00%

Steps to perform Forward Feature Selection

1. Train n model using each feature (n) individually and check the performance
2. Choose the variable which gives the best performance

Variable_Selected
Calories_burnt

Variable used	Accuracy
Calories_burnt	87.00%
Gender	80.00%
Plays_Sport?	85.00%

Steps to perform Forward Feature Selection

1. Train n model using each feature (n) individually and check the performance
2. Choose the variable which gives the best performance
3. Repeat the process and add one variable at a time

Steps to perform Forward Feature Selection

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 Forward Feature Selection

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 Forward Feature Selection

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 Forward Feature Selection

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%

Steps to perform Forward Feature Selection

1. Train n model using each feature (n) individually and check the performance
2. Choose the variable which gives the best performance
3. Repeat the process and add one variable at a time
4. Variable producing the highest improvement is retained

Steps to perform Forward Feature Selection

Variable added	Accuracy
Gender	88.00%
Plays_Sport?	91.00%

Steps to perform Forward Feature Selection

Variable added	Accuracy
Gender	88.00%
Plays_Sport?	91.00%

Steps to perform Forward Feature Selection

Variable_Selected
Calories_burnt

Variable added	Accuracy
Gender	88.00%
Plays_Sport?	91.00%

Steps to perform Forward Feature Selection

Variable_Selected
Calories_burnt
Plays_Sport?

Variable added	Accuracy
Gender	88.00%
Plays_Sport?	91.00%

Steps to perform Forward Feature Selection

1. Train n model using each feature (n) individually and check the performance
2. Choose the variable which gives the best performance
3. Repeat the process and add one variable at a time
4. Variable producing the highest improvement is retained
5. Repeat the entire process until there is no significant improvement in the model’s performance

Thank
You!

Analytics Vidhya
Learn everything about analytics