Feature Selection : Low Variance

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 : Low Variance

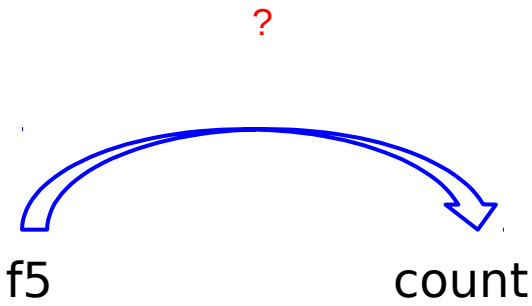
ID	season	holiday	workingday	weather	f5	temp	atemp	humidity	windspeed	count
AB101	1	0	0	1	7	9.84	14.395	81	0.0000	16
AB102	1	0	0	1	7	9.02	13.635	80	0.0000	40
AB103	1	0	0	1	7	9.02	13.635	80	0.0000	32
AB104	1	0	0	1	7	9.84	14.395	75	0.0000	13
AB105	1	0	0	1	7	9.84	14.395	75	0.0000	1
AB106	1	0	0	2	7	9.84	12.880	75	6.0032	1
AB107	1	0	0	1	7	9.02	13.635	80	0.0000	2
AB108	1	0	0	1	7	8.20	12.880	86	0.0000	3
AB109	1	0	0	1	7	9.84	14.395	75	0.0000	8
AB110	1	0	0	1	7	13.12	17.425	76	0.0000	14

Feature Selection : Low Variance

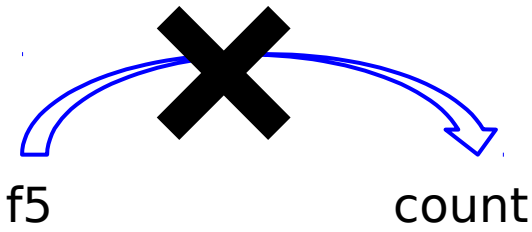
ID	season	holiday	workingday	weather	f5	temp	atemp	humidity	windspeed	count
AB101	1	0	0	1	7	9.84	14.395	81	0.0000	16
AB102	1	0	0	1	7	9.02	13.635	80	0.0000	40
AB103	1	0	0	1	7	9.02	13.635	80	0.0000	32
AB104	1	0	0	1	7	9.84	14.395	75	0.0000	13
AB105	1	0	0	1	7	9.84	14.395	75	0.0000	1
AB106	1	0	0	2	7	9.84	12.880	75	6.0032	1
AB107	1	0	0	1	7	9.02	13.635	80	0.0000	2
AB108	1	0	0	1	7	8.20	12.880	86	0.0000	3
AB109	1	0	0	1	7	9.84	14.395	75	0.0000	8
AB110	1	0	0	1	7	13.12	17.425	76	0.0000	14

Analytics Vidhya
Learn everything about analytics

Feature Selection : Low Variance



Feature Selection : Low Variance



Feature Selection : Low Variance

Variance = 0

Feature Selection : Low Variance

Variance  Spread of the data

Feature Selection : Low Variance

$$\sigma^2 = \frac{\Sigma(x-\bar{x})^2}{n}$$

Feature Selection : Low Variance

Variance = 0

Feature Selection : Low Variance

- Decide a threshold variance value
- Drop the variables having less variance than the threshold
  - Normalize the variables before calculating variance

Thank
You!

Analytics Vidhya
Learn everything about analytics