

Documentação do tratamento do banco de dados após tarefa de clustering hierárquico SOM/K-médias

Metodologias utilizadas:

- I) 294 amostras entre Normal, HI e LI
- II) 294 amostras entre Normal e LI (as amostras de HI passaram a ser do tipo Normal)
- III) 168 amostras entre Normal e LI (amostras HI foram excluídas)

Vetor de amostras:

Os atributos utilizados foram os seguintes: (data0X.data e data0X.data_d)

data0X. data = matriz de dados normalizados

data0X. data_d = matriz de dados denormalizados

1 - 2a Harmônica
2 - 3a Harmônica
3 - 5a Harmônica
4 - 7a Harmônica
5 - 0,5*Harmônica fundamental
6 - 1,5*Harmônica fundamental
7 - Porcentagem de carga aplicada

Os rótulos a serem utilizados para classificação são os seguintes: (data0X.labels)

1 – Tipo da falha (Normal, HI ou LI)
2 – Gravidade da falha (0, 1, 2 ou 3)
3 – Cluster em que foi alocado (entre 1 e 8)

Metodologia I:

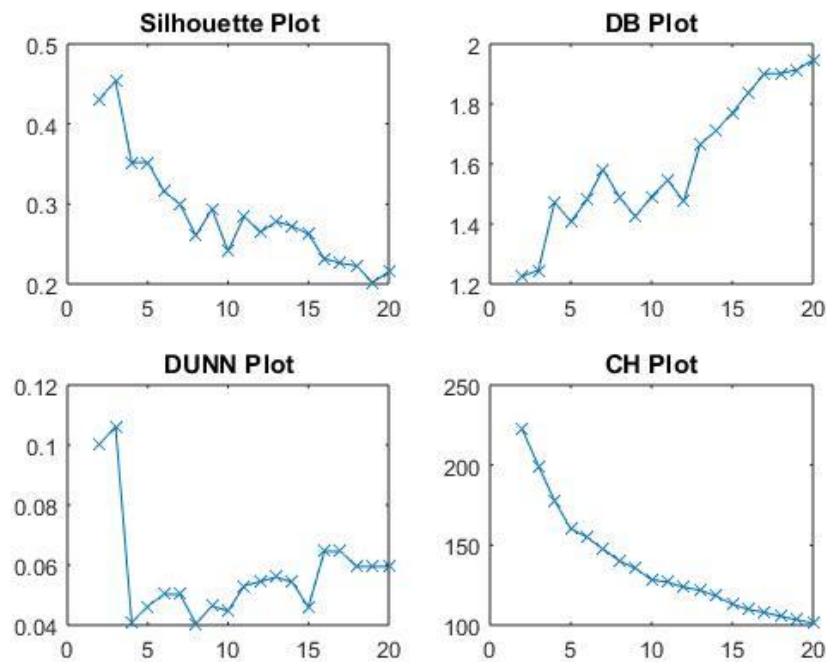
data01 =

data: [7x294 double]

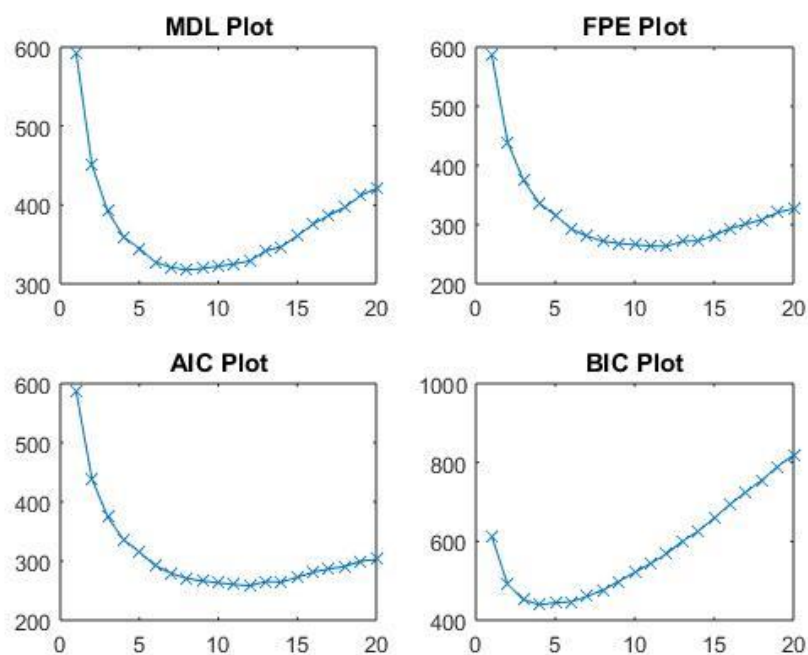
data_d: [7x294 double]

labels: [3x294 double]

Resposta dos índices:



Metodologia I: Índices relativos



Metodologia I: Critérios de informação

Metodologia I:

K=2	Fault Extent			Fault Level				
	Normal	HI	LI	0	1	2	3	
Cluster 1	11	42	43	11	28	30	27	96
Cluster 2	31	84	83	31	56	56	55	198

K=3	Fault Extent			Fault Level				
	Normal	HI	LI	0	1	2	3	
Cluster 1	27	72	71	27	48	48	47	170
Cluster 2	11	42	43	11	28	30	27	96
Cluster 3	4	12	12	4	8	8	8	28

K=4	Fault Extent			Fault Level				
	Normal	HI	LI	0	1	2	3	
Cluster 1	13	39	29	13	26	21	21	81
Cluster 2	4	12	12	4	8	8	8	28
Cluster 3	14	33	42	14	22	27	26	89
Cluster 4	11	42	43	11	28	30	27	96

K=8	Fault Extent			Fault Level				
	Normal	HI	LI	0	1	2	3	
Cluster 1	14	7	5	14	7	3	2	26
Cluster 2	0	0	20	0	0	8	12	20
Cluster 3	9	30	23	9	23	12	18	62
Cluster 4	9	21	20	9	14	15	12	50
Cluster 5	4	12	10	4	8	8	6	26
Cluster 6	0	35	18	0	21	19	13	53
Cluster 7	0	0	16	0	0	2	14	16
Cluster 8	6	21	14	6	11	19	5	41

Metodologia I:

*) K ótimo escolhido K=8

Observações

Cluster 1: Normal em destaque (Amostras de falta removidas)

Clusters 2, 6 e 7: Faltosos exclusivos

Clusters 3, 4, 5 e 8: Faltosos em maioria
(amostras Normal excluídas)

Após remoções:

K=8	Fault Extent			Fault Level				
	Normal	HI	LI	0	1	2	3	
Cluster 1	14	0	0	14	0	0	0	14
Cluster 2	0	0	20	0	0	8	12	20
Cluster 3	0	30	23	0	23	12	18	53
Cluster 4	0	21	20	0	14	15	12	41
Cluster 5	0	12	10	0	8	8	6	22
Cluster 6	0	35	18	0	21	19	13	53
Cluster 7	0	0	16	0	0	2	14	16
Cluster 8	0	21	14	0	11	19	5	35

data01_red =

data: [7x254 double]
data_d: [7x254 double]
labels: [3x254 double]

Metodologia II:

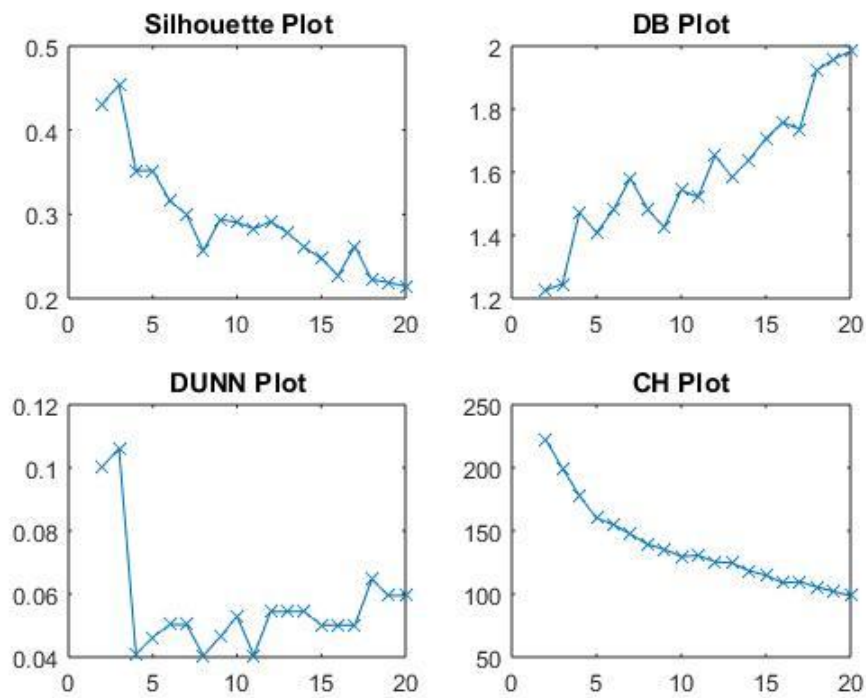
data02 =

data: [7x294 double]

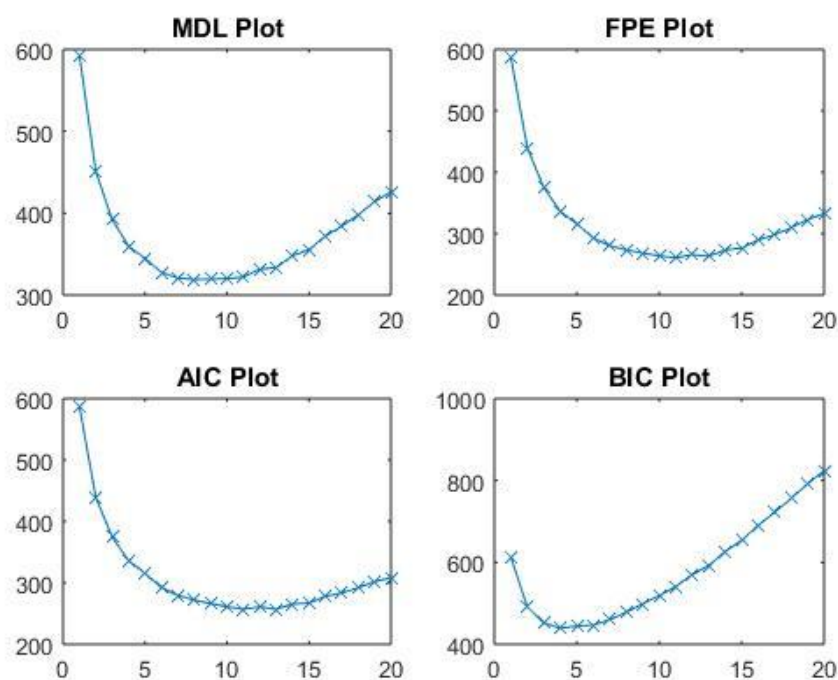
data_d: [7x294 double]

labels: [3x294 double]

Resposta dos índices:



Metodologia II: Índices relativos



Metodologia II: Critérios de informação

Metodologia II:

K=2	Fault Extent		Fault Level				Total
	Normal	LI	0	1	2	3	
Cluster 1	53	43	53	14	16	13	96
Cluster 2	115	83	115	28	28	27	198

K=3	Fault Extent		Fault Level				Total
	Normal	LI	0	1	2	3	
Cluster 1	53	43	53	14	16	13	96
Cluster 2	99	71	99	24	24	23	170
Cluster 3	16	12	16	4	4	4	28

K=4	Fault Extent		Fault Level				Total
	Normal	LI	0	1	2	3	
Cluster 1	52	29	52	13	9	7	81
Cluster 2	16	12	16	4	4	4	28
Cluster 3	53	43	53	14	16	13	96
Cluster 4	47	42	47	11	15	16	89

K=8	Fault Extent		Fault Level				Total
	Normal	LI	0	1	2	3	
Cluster 1	42	23	42	11	5	7	65
Cluster 2	35	18	35	11	7	0	53
Cluster 3	16	10	16	4	4	2	26
Cluster 4	21	5	21	3	1	1	26
Cluster 5	24	14	24	6	8	0	38
Cluster 6	0	16	0	0	2	14	16
Cluster 7	30	20	30	7	9	4	50
Cluster 8	0	20	0	0	8	12	20

Metodologia II:

*) K ótimo escolhido K=8

Observações

Cluster 4: Predominantemente Normal (5 amostras LI removidas)

Clusters 6 e 8: LI exclusivos

Clusters 3 e 7: Aproximadamente 1 Normal para 0,5 LI
(cluster excluído, Normal considerado não predominante)

Clusters 1, 2 e 5: Aproximadamente 2 Normal para 1 LI
(23 amostras LI de Cluster 1, 18 amostras de Cluster 2 e 14 amostras de Cluster 5 excluídas)
(Considerou-se cluster predominantemente Normal)

Após remoções:

K=8	Fault Extent		Fault Level				Total
	Normal	LI	0	1	2	3	
Cluster 1	42	0	42	0	0	0	42
Cluster 2	35	0	35	0	0	0	35
Cluster 3	0	0	0	0	0	0	0
Cluster 4	21	0	21	0	0	0	21
Cluster 5	24	0	24	0	0	0	24
Cluster 6	0	16	0	0	2	14	16
Cluster 7	0	0	0	0	0	0	0
Cluster 8	0	20	0	0	8	12	20

```
data02_red =
```

```
data: [7x158 double]  
data_d: [7x158 double]  
labels: [3x158 double]
```

Metodologia III:

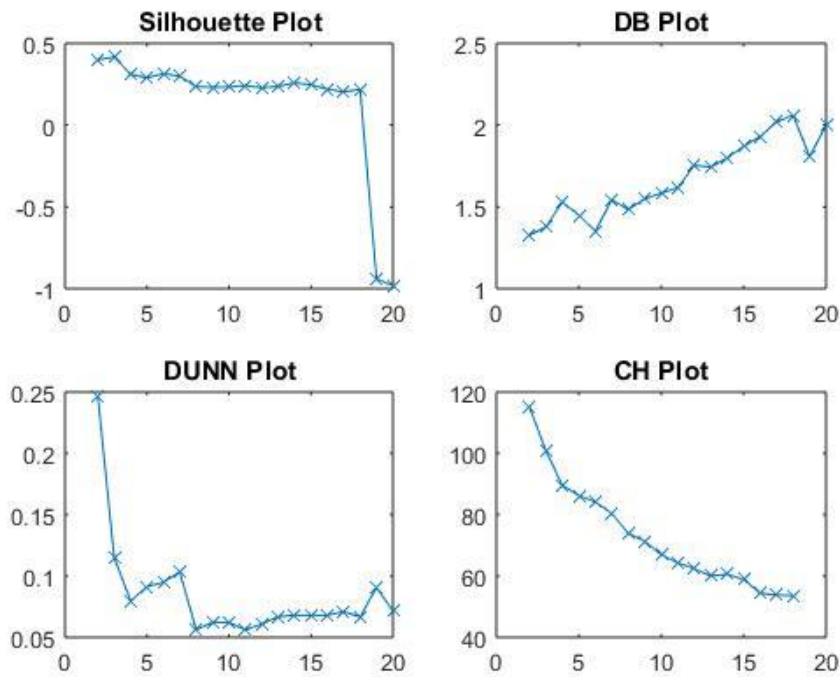
data03 =

data: [7x168 double]

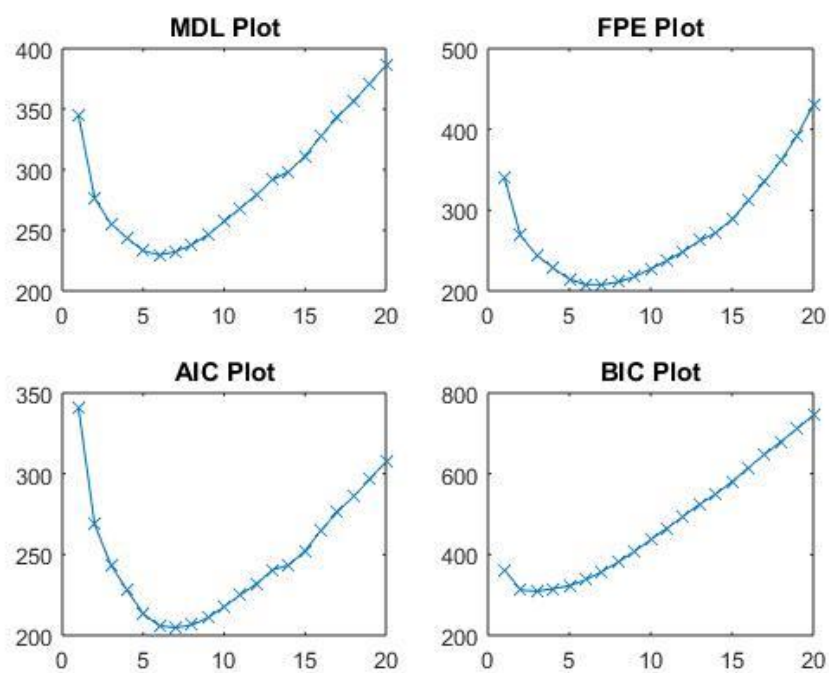
data_d: [7x168 double]

labels: [3x168 double]

Resposta dos índices:



Metodologia III: Índices relativos



Metodologia III: Critérios de informação

Metodologia III:

K=2	Fault Extent		Fault Level				Total
	Normal	LI	0	1	2	3	
Cluster 1	14	43	14	14	16	13	57
Cluster 2	28	83	28	28	28	27	111

K=3	Fault Extent		Fault Level				Total
	Normal	LI	0	1	2	3	
Cluster 1	11	43	11	14	16	13	54
Cluster 2	27	71	27	24	24	23	98
Cluster 3	4	12	4	4	4	4	16

K=6	Fault Extent		Fault Level				Total
	Normal	LI	0	1	2	3	
Cluster 1	0	38	0	12	15	11	38
Cluster 2	15	4	15	2	1	1	19
Cluster 3	9	25	9	11	7	7	34
Cluster 4	4	12	4	4	4	4	16
Cluster 5	14	38	14	13	16	9	52
Cluster 6	0	9	0	0	1	8	9

K=7	Fault Extent		Fault Level				Total
	Normal	LI	0	1	2	3	
Cluster 1	0	19	0	0	7	12	19
Cluster 2	4	12	4	4	4	4	16
Cluster 3	15	4	15	2	1	1	19
Cluster 4	9	25	9	11	7	7	34
Cluster 5	14	37	14	13	16	8	51
Cluster 6	0	21	0	12	9	0	21
Cluster 7	0	8	0	0	0	8	8

K=8	Fault Extent		Fault Level				Total
	Normal	LI	0	1	2	3	
Cluster 1	14	3	14	2	0	1	17
Cluster 2	9	20	9	9	4	7	29
Cluster 3	0	19	0	0	7	12	19
Cluster 4	0	21	0	12	9	0	21
Cluster 5	4	12	4	4	4	4	16
Cluster 6	0	13	0	0	1	12	13
Cluster 7	6	13	6	5	8	0	19
Cluster 8	9	25	9	10	11	4	34

Metodologia III:

*) K ótimo escolhido K=8

Observações

Cluster 1: Predominantemente Normal (3 amostras LI removidas)

Clusters 3, 4 e 6: LI exclusivos

Clusters 2 e 7: Aproximadamente 1 Normal para 2 LI
(cluster excluído, LI considerado não predominante)

Clusters 5 e 8: Aproximadamente 1 Normal para 3 LI
(4 amostras Normal de Cluster 5 e 9 amostras de Cluster 8 excluídas)
(Considerou-se cluster predominantemente LI)

Após remoções:

K=8	Fault Extent		Fault Level				Total
	Normal	LI	0	1	2	3	
Cluster 1	14	3	14	2	0	1	17
Cluster 2	0	0	0	0	0	0	0
Cluster 3	0	19	0	0	7	12	19
Cluster 4	0	21	0	12	9	0	21
Cluster 5	0	12	0	4	4	4	12
Cluster 6	0	13	0	0	1	12	13
Cluster 7	0	0	0	0	0	0	0
Cluster 8	0	25	0	10	11	4	25

data03_red =

data: [7x104 double]
data_d: [7x104 double]
labels: [3x104 double]