

Redes móveis LTE – Falhas e degradações

Diagnóstico de causa raiz de falhas e degradação de indicadores de performance em redes LTE utilizando agrupamento não supervisionado e classificação supervisionada

CPE 722 – Redes Neurais Não Supervisionadas e Agrupamento

Prof. Luiz Pereira Calôba, Dr.Ing.

Carlos Eduardo Covas Costa

2017 / 1

Agenda

1

Introdução

2

Descrição do Dataset

3

Aprendizado não supervisionado

4

Aprendizado Supervisionado

5

Avaliação do resultados

6

Conclusão

Motivação

- Crescimento constante no tamanho e complexidade de redes móveis celulares LTE.
- Dificuldade maior para realização de tarefas de monitoramento e diagnóstico de falhas a medida que a quantidade de elementos de rede aumenta.
- Interesse crescente na capacidade de ‘Self-Healing’ de uma rede detectar suas falhas, evitando que problemas como queda de chamada, cobertura ruim e baixa velocidade permaneçam bastante tempo sem serem percebidos.
- Sistemas de diagnóstico de falhas normalmente utilizam regras definidas a priori (IF-ELSE) ou se baseiam em bases com rótulos de problemas de redes que dificilmente são produzidas.

Objetivo

- Utilizar técnica de aprendizado não supervisionado para detecção de padrões de falhas de rede LTE.
- Identificar os padrões detectados e perceber e atribuir rótulos para classes baseado no comportamento estatístico dos dados e em conhecimento específico da área.
- Comparar o a técnica não supervisionado com uma técnica supervisionada utilizando o mesmo dataset .

Dataset

Dataset gerado utilizando um simulador de redes LTE implementado em MATLAB [2].

57 células distribuídas igualmente em uma grade hexagonal

Dataset disponível em [3]

Parameter	Configuration
Cellular layout	Hexagonal grid, 57 cells, cell radius 0.5 km
Transmission direction	Downlink
Carrier frequency	2.0 GHz
System bandwidth	1.4 MHz (6 physical resource blocks)
Frequency reuse	1
Propagation model	Okumura-Hata with wrap-around, Log-normal slow fading, $\sigma_{sf} = 8$ dB and correlation distance=50m
Channel model	Multipath fading, ETU model
Mobility model	Random direction, 3 km/h
Service model	Full Buffer, poisson traffic arrival
Base station model	Tri-sectorized antenna, SISO, Donwtilt=9 ° $P_{TX_{max}} = 43$ dBm, Azimuth beamwidth (AB)=70 ° Elevation beamwidth (EB)=10 °
Scheduler	Time domain: Round-Robin, Frequency domain: Best Channel
Power control	Equal transmit power per physical resource blocks
Link Adaptation	Fast, CQI based, perfect estimation
Handover	Triggering event = A3, HOM = 3 dB, Measurement type = RSRP
Radio Link Failure	SINR < -6.9 dB for 500 ms
Traffic distribution	Evenly distributed in space
Time resolution	100 TTI (100 ms)
Epoch & KPI time	100 s

Entradas composta por indicadores KPIs 3GPP[4] de um célula em um determinado instante de tempo.

$S = [\text{Retainability}, \text{HOSR}, \text{RSRP}, \text{RSRQ}, \text{SINR}, \text{Throughput}, \text{Distance}]$

Retainability: Conexões terminadas com sucesso dividido pelo total de conexões estabelecidas.

HOSR: Taxa de sucesso de Handover (mobilidade)

RSRP: Reference Signal Received Power. Potência média dos sinais de referência da largura de banda usada.

RSRQ: Reference Signal Received Quality. Razão da potência útil (RSRP) e toda potência recebida (RSSI).

SINR: Relação sinal ruído.

Throughput: Média da velocidade dos usuários na célula.

Distance: Distância entre a estação e o usuário

Dataset

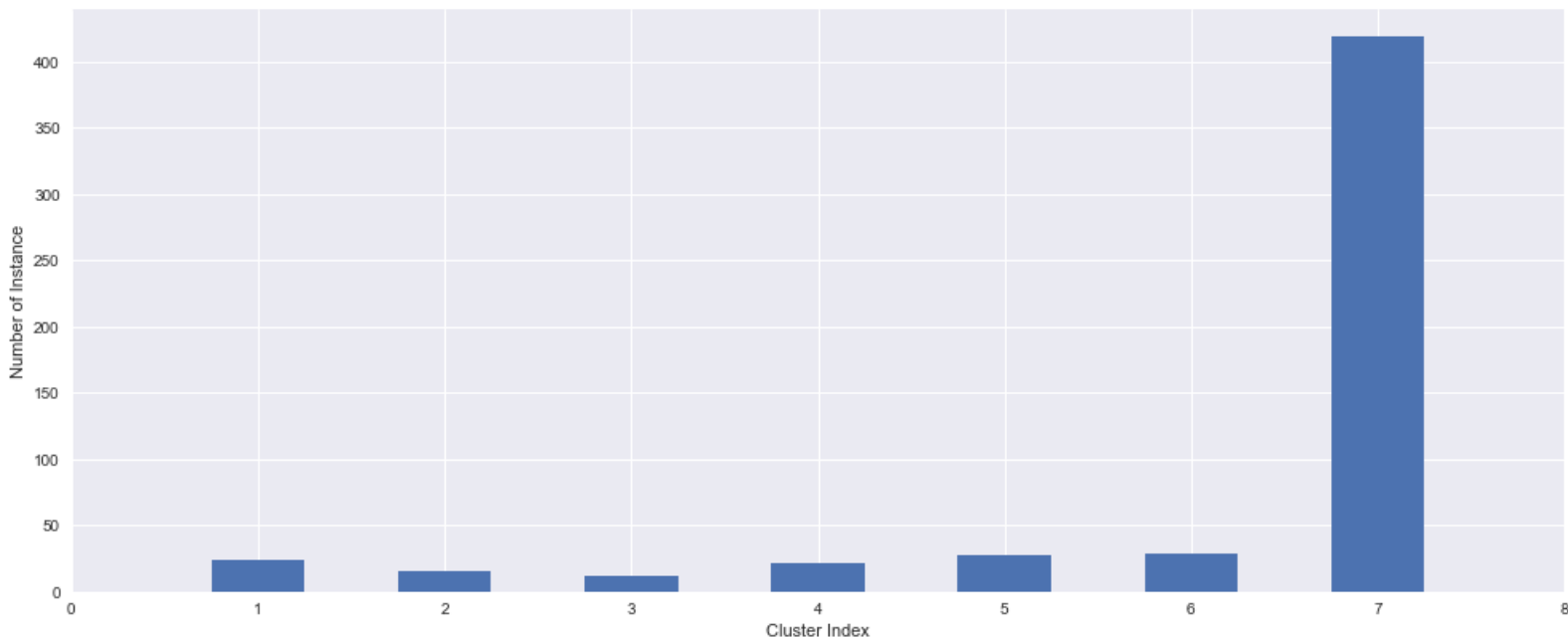
6 diferentes tipos de falhas foram simulados para deteriorar células escolhidas aleatoriamente.

Células em estado normal são mais comuns que células com falhas em uma rede. Neste trabalho a proporção de células em estado normal é de 73.93%

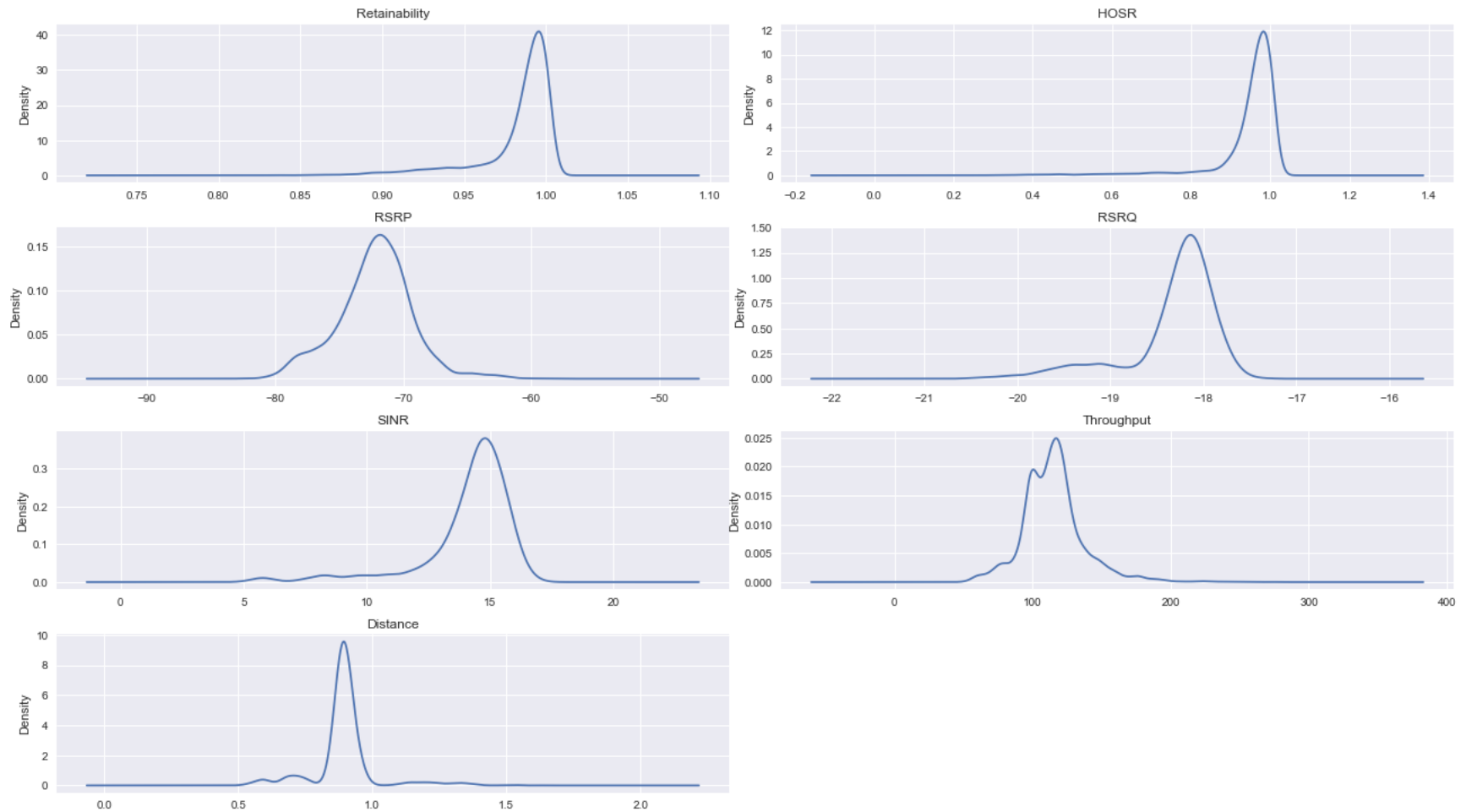
Treinamento: 550 registros

Validação: 4009

Fault Cause	Configuration	Number of Cases	
		Training	Validation
EU	Downtilt=[0,1] °	32	212
ED	Downtilt=[16,15,14] °	28	212
ERP	$\Delta P_{TX} = [7,8,9,10]$ dB	28	208
CH	$\Delta_{hole} = [49,50,52,53]$ dBm	14	103
TLHO	HOM=[6,7,8] dBm	34	204
II	$P_{TX_{max}} = 33$ dBm		
	Downtilt=15 °		
	AB=[30, 60] °	15	106
	EB=10 °		
No fault	Normal	399	2964

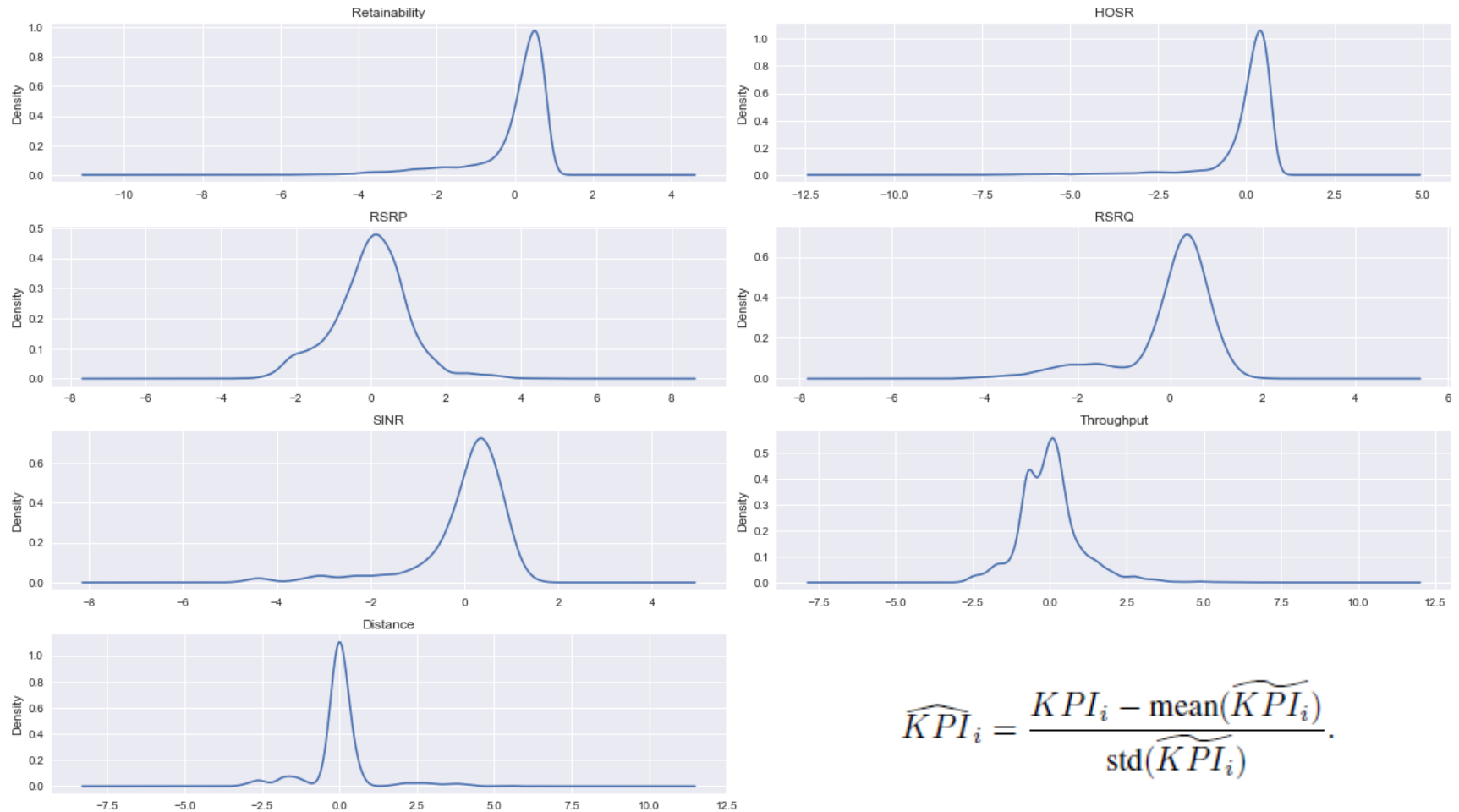


Dataset



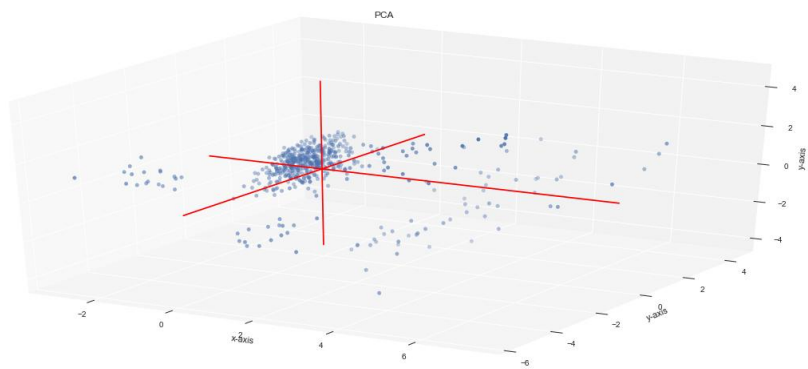
Dataset

Todos os KPIs foram normalizados com média 0 e desvio padrão 1



$$\widehat{KPI}_i = \frac{KPI_i - \text{mean}(\widehat{KPI}_i)}{\text{std}(\widehat{KPI}_i)}.$$

Aprendizado Não Supervisionado (SOM)

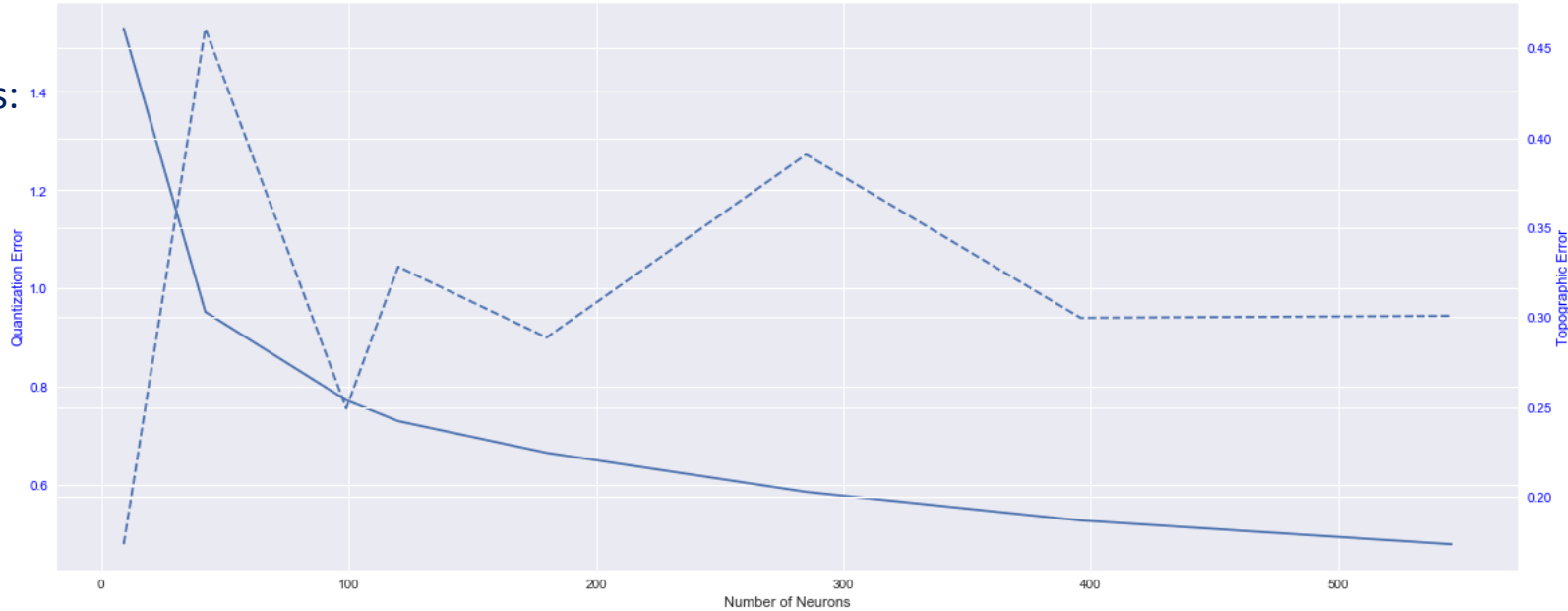


Parâmetros	Valores
Dimensionamento P/Q	$\frac{\sigma^2 da 1^a Componente principal}{\sigma^2 da 2^a Componente principal}$
Treinamento	Batch
Inicialização	PCA
Iterações	20000
Função de Vizinhaça	Bubble
Raio de Vizinhaça Inicial	$.2\sqrt{P^2 + Q^2} = 2.84$
Raio de Vizinhaça Final	1
Decréscimo do Raio	Linear
Taxa de aprendizado inicial	0.1
Decréscimo da taxa de aprendizado	Linear

P/Q = 1.27

Valores Testados:

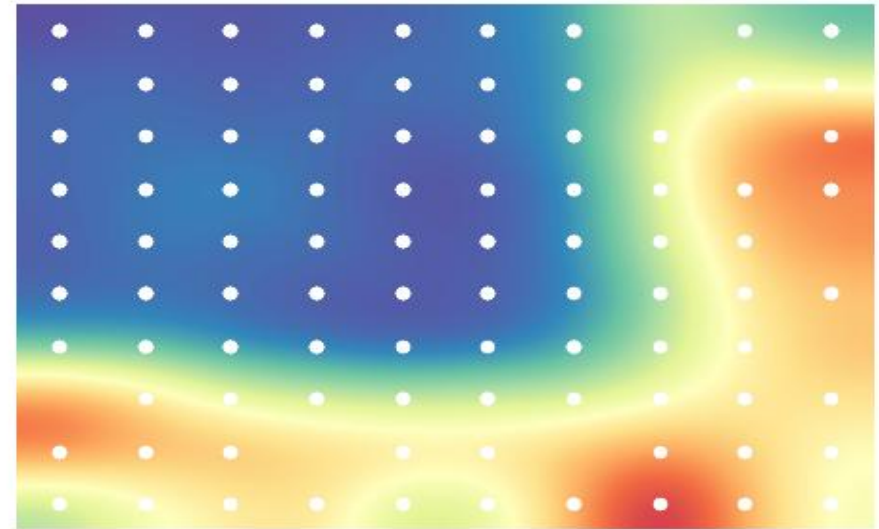
- 3x3
- 6x7
- 9x11
- 10x12
- 12x15
- 15x19
- 18x22
- 21x26



Aprendizado Não Supervisionado

1. Treinamento não supervisionado utilizando rede SOM.
2. Avaliação do número de agrupamentos utilizando Silhouette Index
3. Clusterização hierárquica não supervisionada utilizando método Ward.
4. Análise das pdfs de cada KPI para cada classe encontrada para identificação do tipo de falha.

U-Matrix



Mapa de Kohonen calculado utilizando biblioteca Somoclu de Python[5]

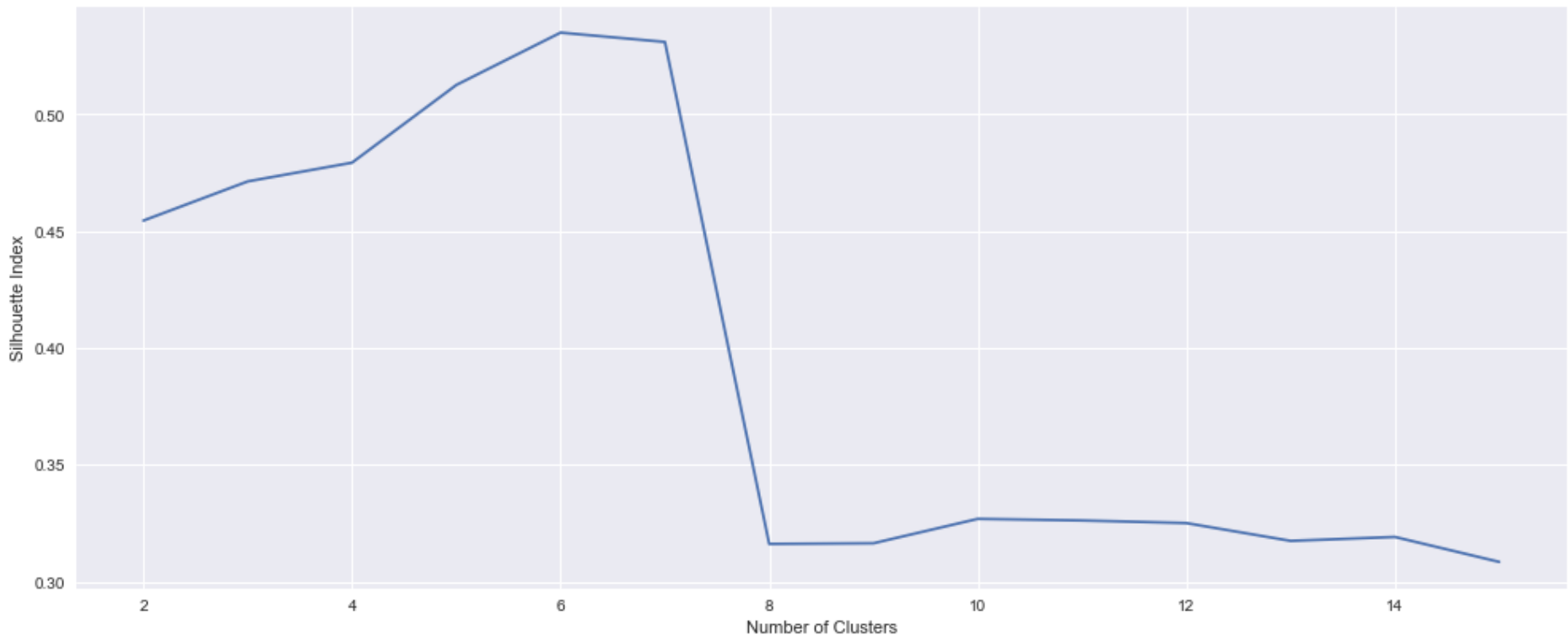
Aprendizado Não Supervisionado

Silhouette Coefficient:

$$s = \frac{b - a}{\max(a, b)}$$

a é distância média entre uma amostra e todos os outros pontos da mesma classe.

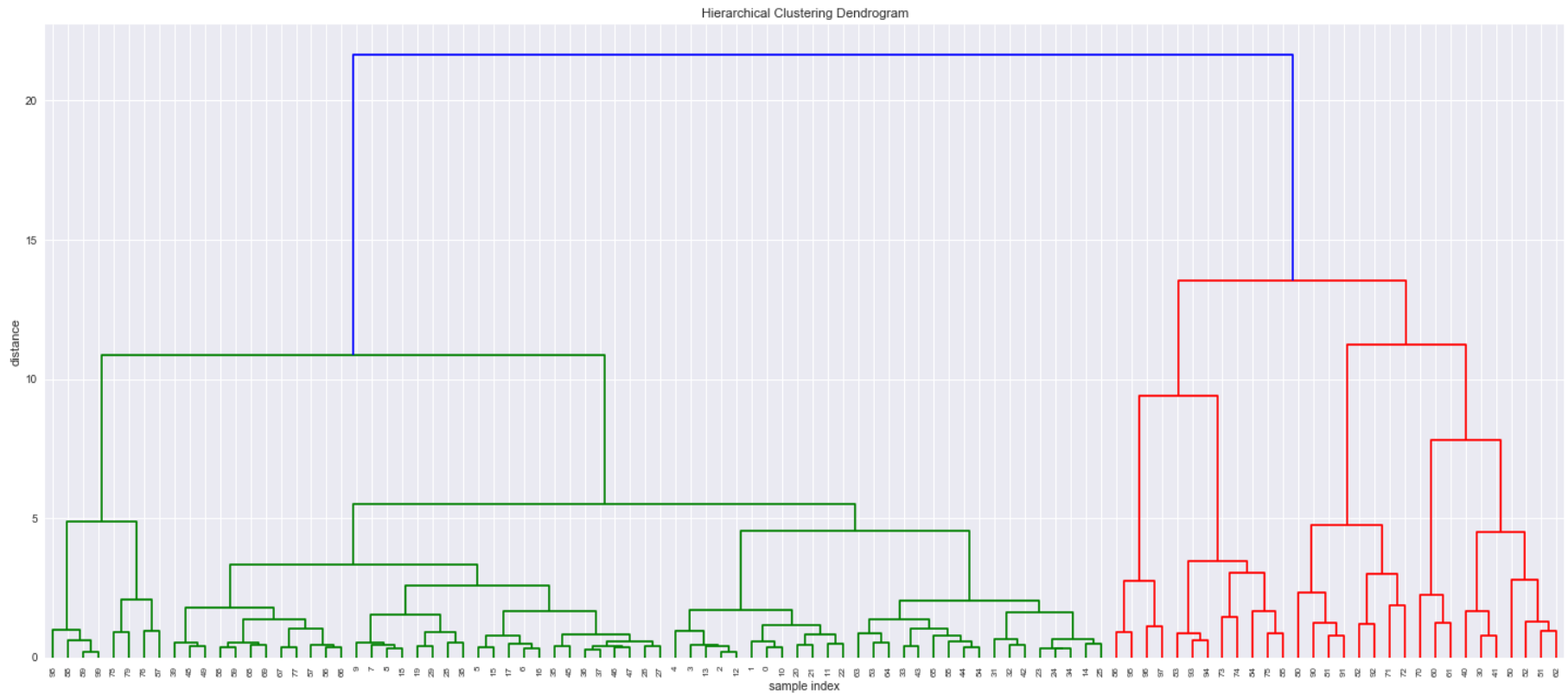
b é distância média entre uma amostra e todos os outros pontos no próximo cluster mais próximo.



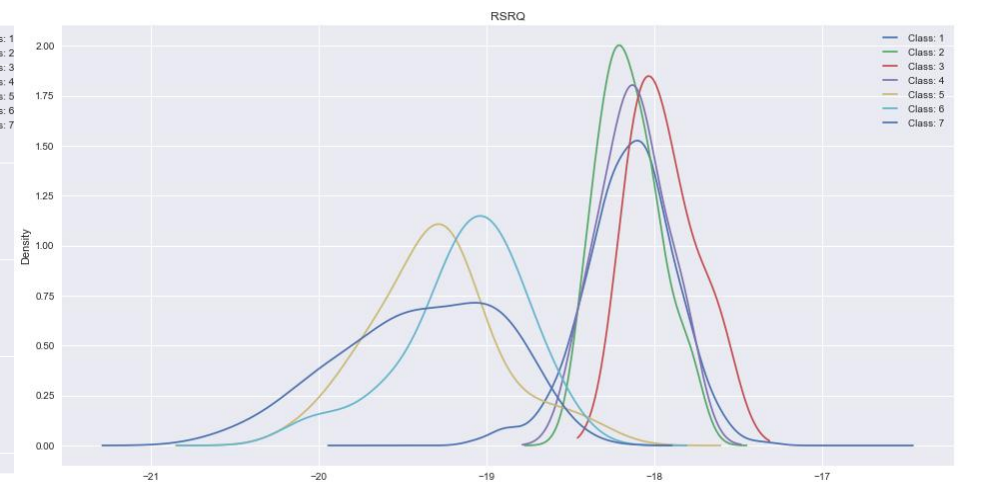
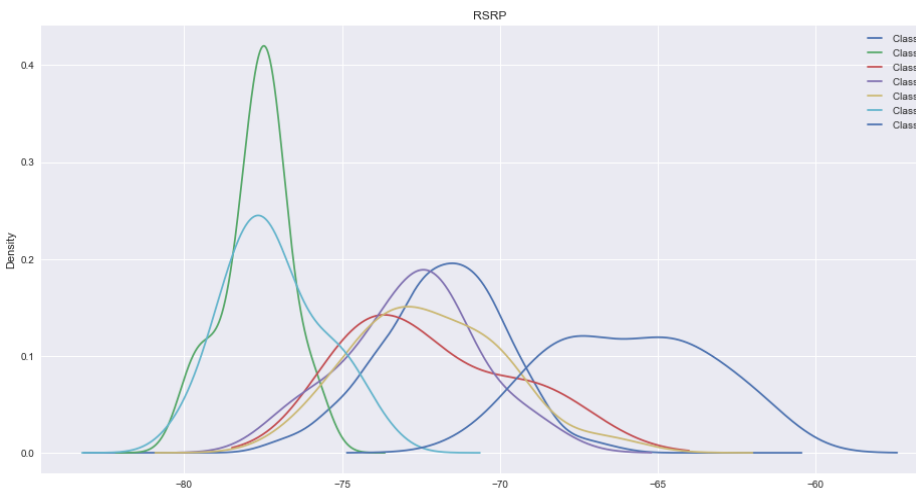
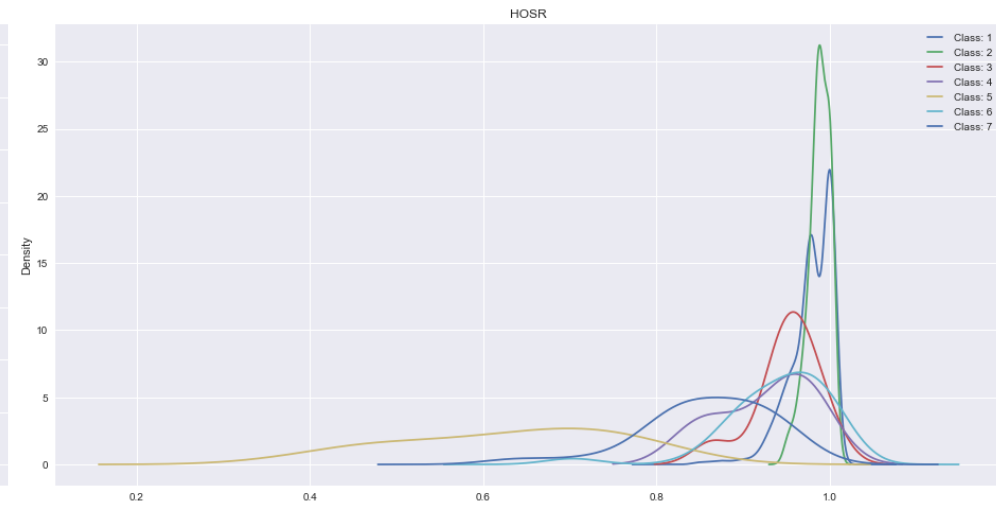
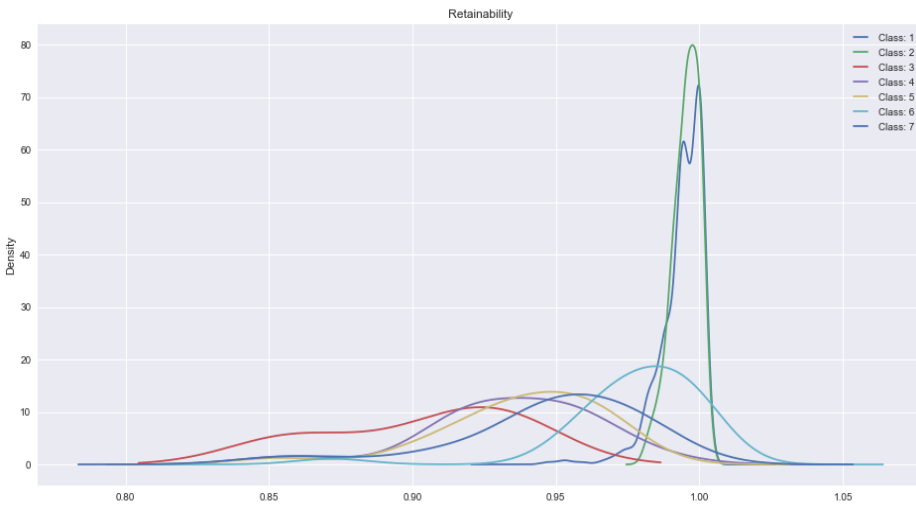
Aprendizado Não Supervisionado

Clusterização aglomerativa hierárquica baseada no critério Ward (método de mínima variância) como função objetivo.

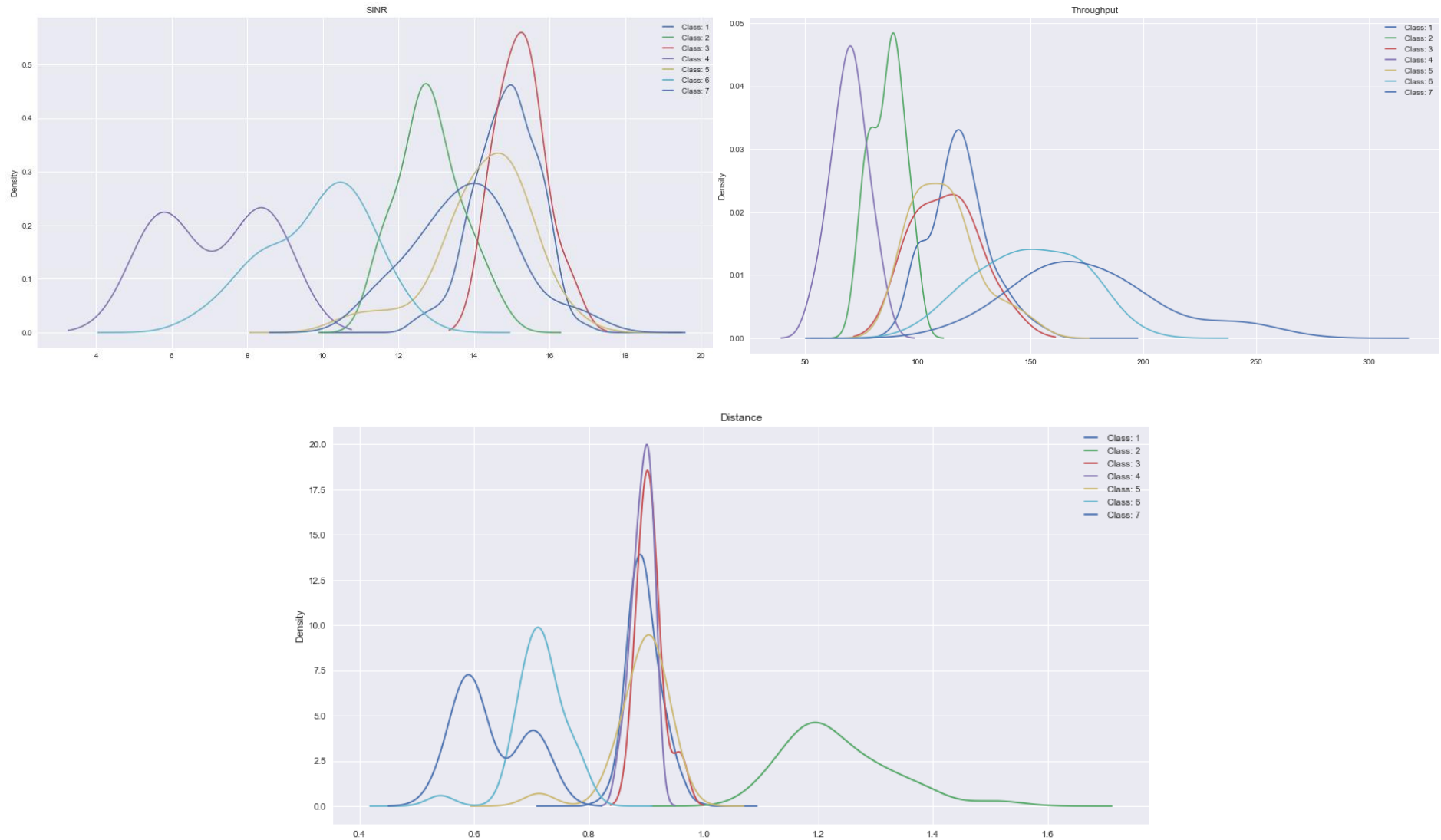
Escolhido $K = 7$ clusters



Pdf (KPIs por causa)



Pdf (KPIs por causa)



Identificação de causas (classes)

	Retainability	HOSR	RSRP	RSRQ	SINR	Throughput	Distance
class							
1	0.993707	0.978094	-71.821918	-18.148264	14.847058	118.550412	0.896384
2	0.995344	0.987531	-77.703219	-18.150281	12.831062	86.665656	1.234625
3	0.903300	0.950600	-72.182900	-17.943500	15.239200	112.678000	0.907000
4	0.934333	0.926467	-72.803600	-18.126867	7.105533	69.745067	0.892867
5	0.937952	0.629810	-72.219000	-19.306857	14.257952	111.963048	0.891143
6	0.977862	0.939759	-77.155310	-19.134207	9.844345	149.814483	0.713414
7	0.947538	0.858192	-65.857346	-19.391385	13.774462	176.092385	0.628423

Classe 1 :

Normal

Classe 2  Dist  Throughput

Uptilt excessivo

Classe 3  Retainability

Buraco de cobertura

Classe 4  SINR  Ret  Throughput

Interferência externa

Classe 5  HOSR  RSRQ

Problema de mobilidade (Handover)

Classe 6  Ret  Dist  RSRP  Thp  RSRQ

Potência de Tx reduzida

Classe 7  Ret  Dist  RSRP  Thp  RSRQ

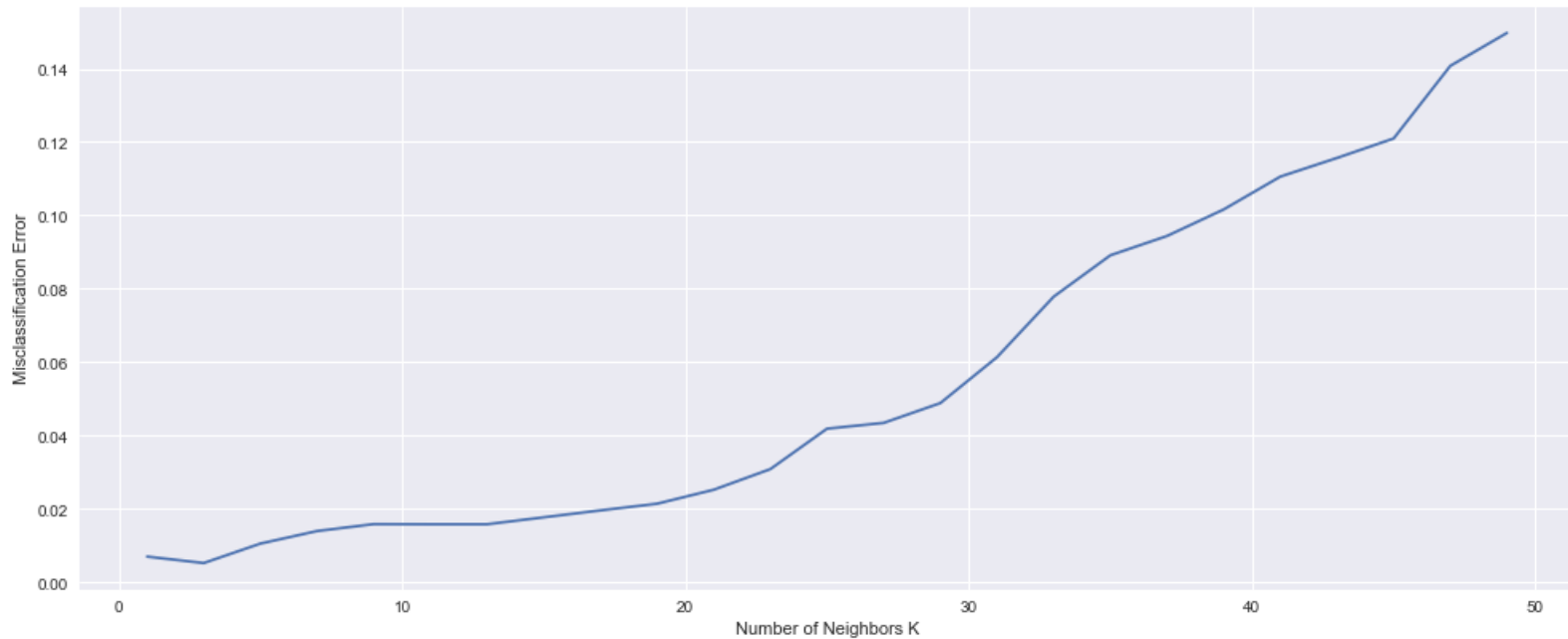
Downtilt excessivo

Aprendizado Supervisionado

Algoritmo: k -nearest neighbors

K com maior acurácia = 3

Validação cruzada 10 fold



Avaliação

Taxa de falso positivo (TFP): Casos normais diagnosticados como falhas pelo número de casos normais

Taxa de falso negativo(TFN): Casos problematicos diagnosticados como normais pelo número de casos problematicos.

Erro de diagnóstico (ED): Casos problematicos diagnosticados com uma falha diferente da real

Prevalência de casos normais(PN): 73.93%

Prevalência de falhas(PF): 26.07%

Erro Total: $PN * TFP + PF * (TFN + ED)$

Fase	Algoritmo	Acurácia[%]	TFP[%]	TFN[%]	ED[%]	Ettotal[%]
Treinamento	KNN	97.19	0	5.3	0.66	1.64
Treinamento	SOM	96.36	0	11.92	1.32	3.63
Teste	KNN	99.57	0.26	1.14	0.76	0.69
Teste	SOM	98.21	0.067	5.16	1.53	1.79

Conclusão

- A grande prevalência de dados pertencente a uma mesma classe (células em estado normal) tornou a base desbalanceada, atrapalhando, assim, uma melhor visualização do mapa de Kohonen
- Apesar do desbalanceamento de classes, o algoritmo SOM+ clusterização hierárquica aglomerativa com método Ward apresentou ótimos resultados identificando as classes normais e falhas produzidas nos dados simulados.
- O método de avaliação dos agrupamentos, baseado em similaridade intra cluster e dissimilaridade inter cluster com o agrupamento mais próximo, encontrou resultado ótimo para $k=7$ classes, como esperado.
- Ambos os métodos tiveram taxa de falso positivo = 0 no treinamento e um valor bem pequeno no teste, mostrando que não tiveram problemas para classificar os casos em que não há falhas na rede.
- Ambos os métodos tiveram uma taxa um pouco elevada de falso negativos. O SOM apresentou cerca de 11% de classificação de normalidade quando há falhas. Tal fato merece ser revisto, pois possui implicações negativas na aplicação em questão.
- Ambos os métodos supervisionado e não supervisionado obtiveram ótimos resultados de acurácia e erro total de classificação de classes no geral. Apesar de obter um resultado um pouco inferior ao método supervisionado, o método não supervisionado obteve resultados satisfatórios, demonstrando que a solução proposta é válida para aplicação de acordo com os objetivos definidos.

Referências

- [1] A. Gómez-Andrades, P. Muñoz, I. Serrano, and R. Barco, “Automatic root cause analysis for LTE networks based on unsupervised techniques,” *IEEE Trans. Veh. Technol.*, vol. 65, no. 4, pp. 2369–2386, Apr. 2016
- [2] P. Muñoz *et al.*, “Computationally efficient design of a dynamic systemlevel LTE simulator,” *Int. J. Electron. Telecommun.*, vol. 57, no. 3, pp. 347–358, Sep. 2011.
- [3] A. Gómez-Andrades *et al.*, “Labelled cases of LTE problems,” 2014.
[Online]. Available: <http://webpersonal.uma.es/de/rbarco/>
- [4] “Physical layer; Measurements,” Third-Generation Partnership Project, Sophia Antipolis Cedex, France, 3GPP TS 25.215.
- [5] Peter Wittek, Shi Chao Gao, Ik Soo Lim, Li Zhao (2015). Somoclu: An Efficient Parallel Library for Self-Organizing Maps.
[Online] Available: <https://arxiv.org/abs/1305.1422>