



FORMAÇÃO INTELIGÊNCIA ARTIFICIAL E MACHINE LEARNING

MACHINE LEARNING – ESTUDANDO ALGORITMOS

K-MEANS

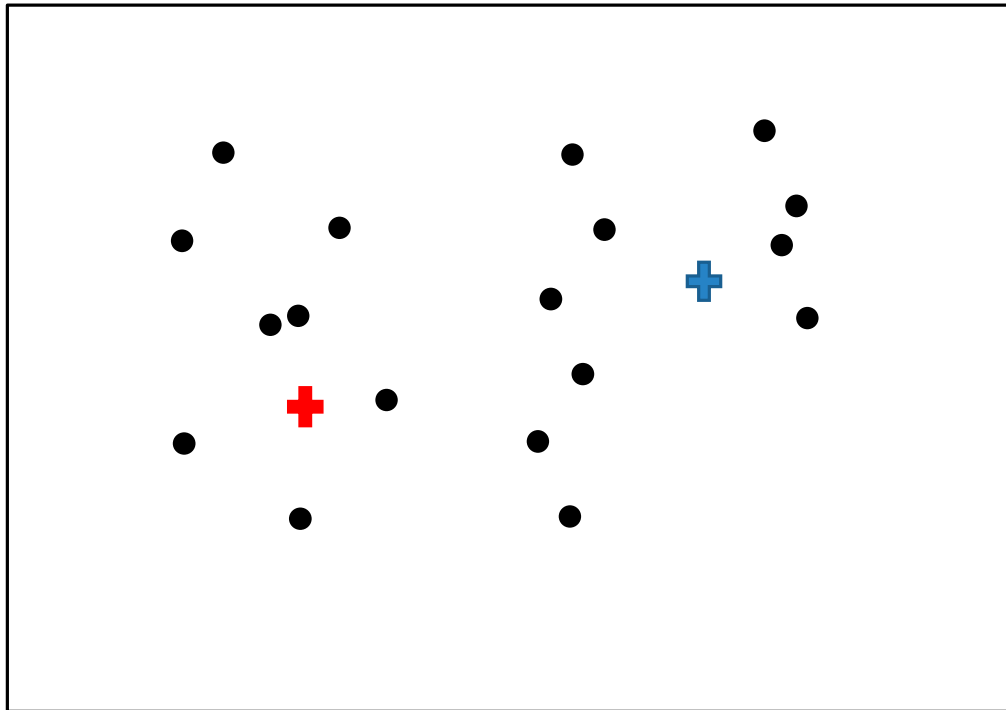
Prof. Fernando Amaral – Todos os Direitos Reservados

K-means

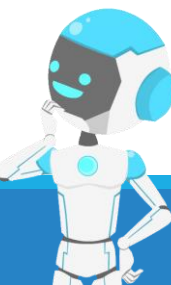
- Dados sem classe
- Agrupamento de forma orgânica (de acordo com algoritmo e seus parâmetros)
- Objetivo:
 - Centroides
 - Grupo para cada instância (rótulo)



Algoritmo de Agrupamento



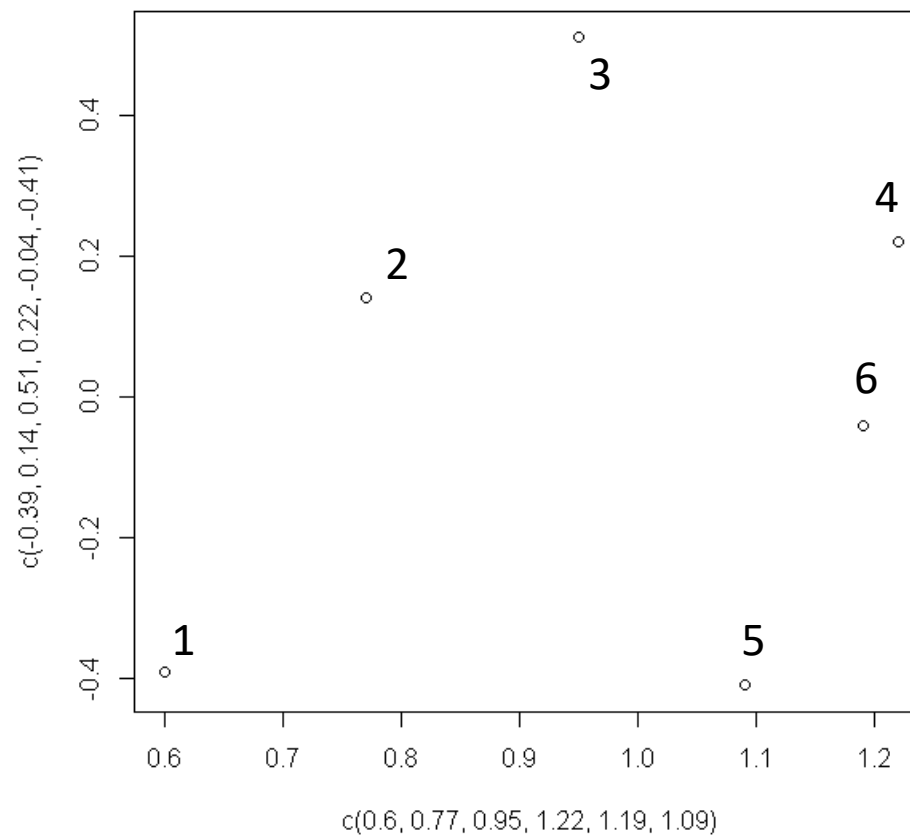
- Definir $K = 2$
- Definir dois centroides
- Cada Instancia busca o centroide mais próximo
- Os centroides são atualizados baseados na media dos pontos do seu grupo
- O processo é repetido até um critério de parada



Instâncias a Agrupar

ID	AtributoA	AtributoB
1	0,6	-0,39
2	0,77	0,14
3	0,95	0,51
4	1,22	0,22
5	1,19	-0,04
6	1,09	-0,41

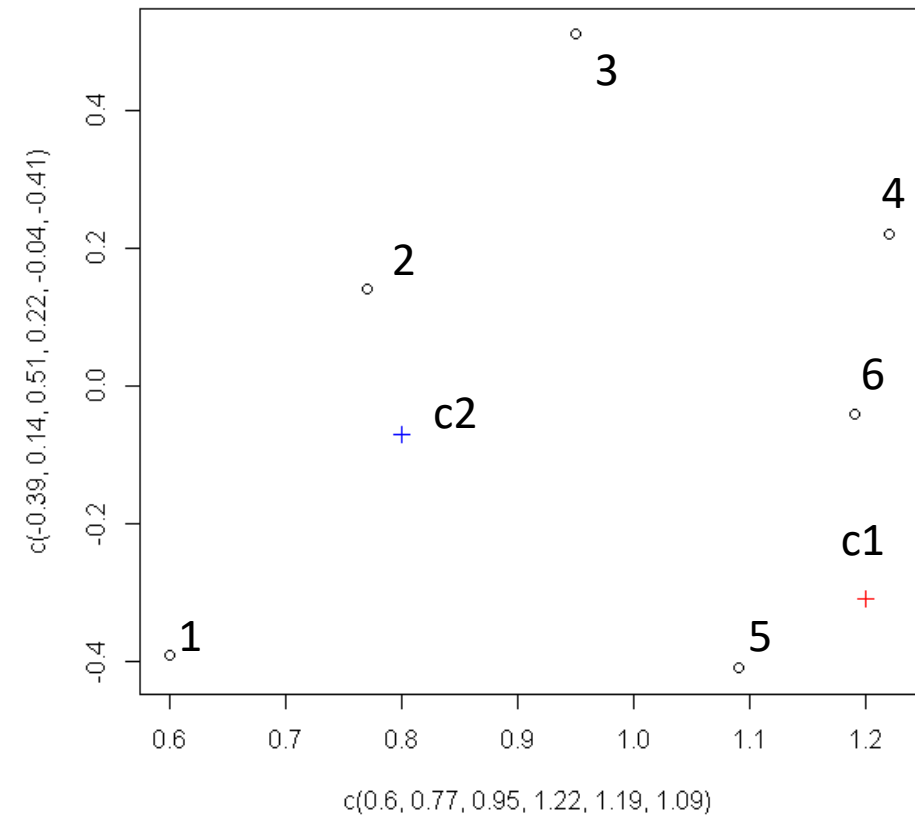
$$\sqrt{(p_x - q_x)^2 + (p_y - q_y)^2}$$



Definição de dois Centroides

ID	AtributoA	AtributoB
1	0,6	-0,39
2	0,77	0,14
3	0,95	0,51
4	1,22	0,22
5	1,19	-0,04
6	1,09	-0,41

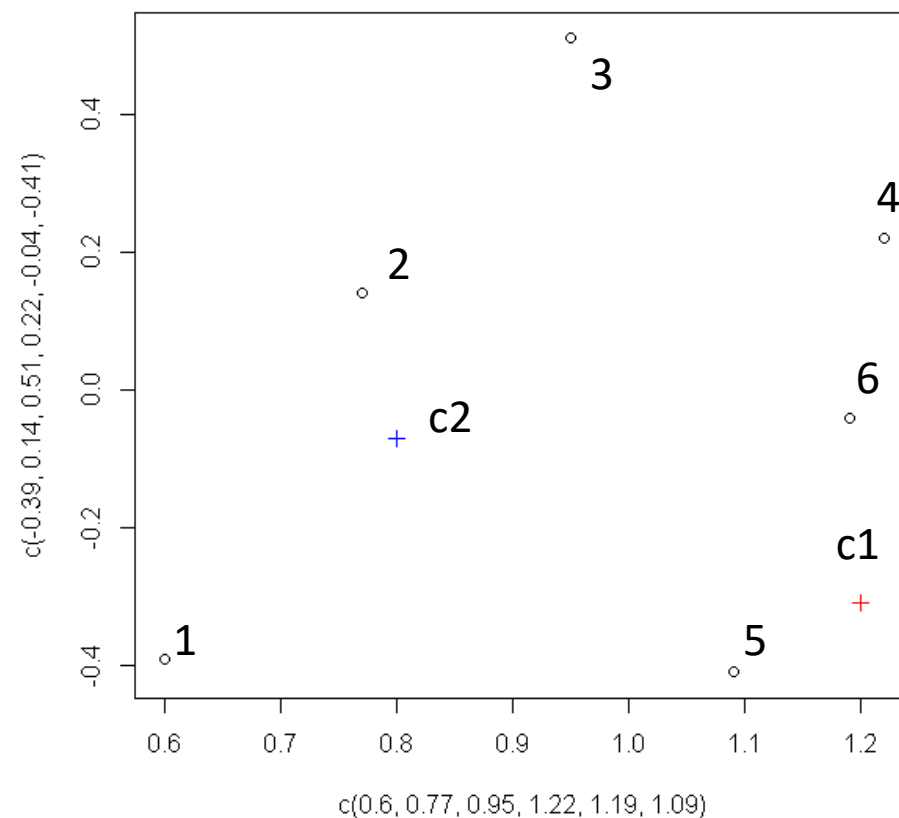
ID	x	y
C1	1,2	-0,31
C2	0,8	-0,07



Cálculo da Distância

ID	x	y
C1	1,2	-0,31
C2	0,8	-0,07

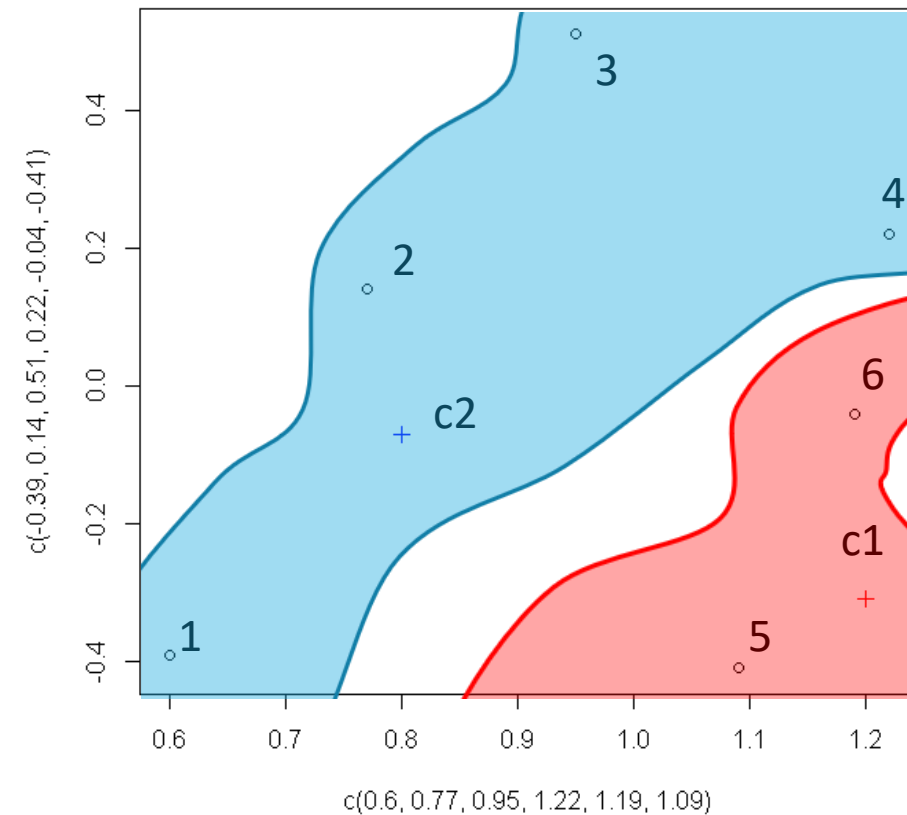
ID	AtributoA	AtributoB	C1	C2
1	0,6	-0,39	0,60	0,37
2	0,77	0,14	0,62	0,21
3	0,95	0,51	0,85	0,59
4	1,22	0,22	0,53	0,51
5	1,19	-0,04	0,27	0,39
6	1,09	-0,41	0,14	0,44



Atribuição ao Grupo

ID	x	y
C1	1,2	-0,31
C2	0,8	-0,07

ID	Atribu toA	Atribu toB	C1	C2
1	0,6	-0,39	0,60	0,37
2	0,77	0,14	0,62	0,21
3	0,95	0,51	0,85	0,59
4	1,22	0,22	0,53	0,51
5	1,19	-0,04	0,27	0,39
6	1,09	-0,41	0,14	0,44

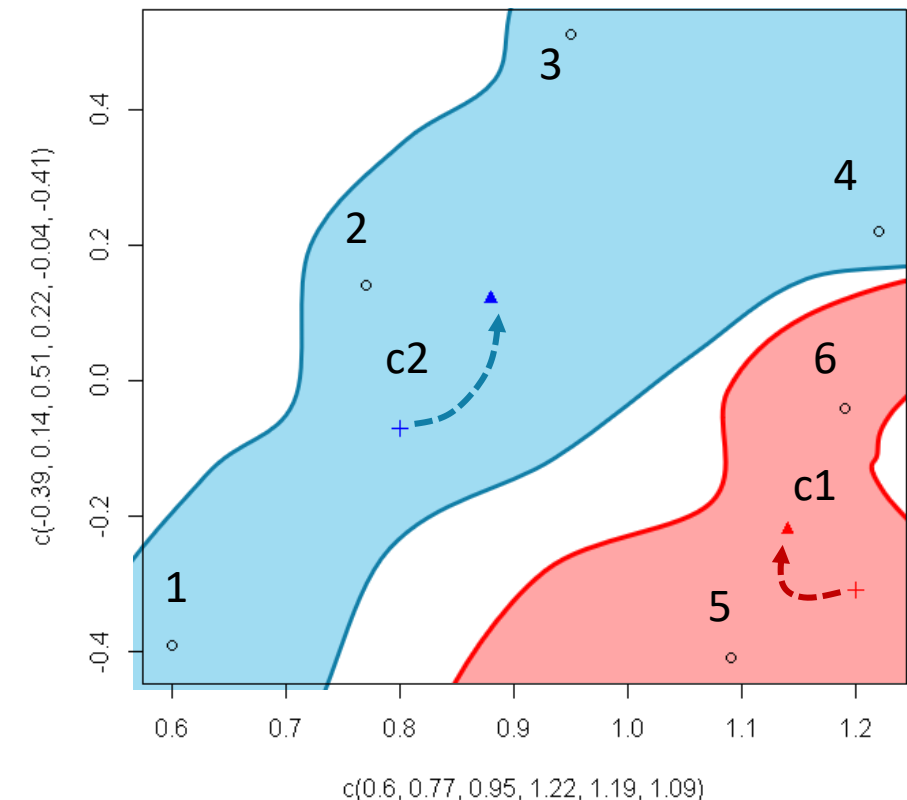


Atualização dos Centroides

Antigo		
ID	x	y
C1	1,2	-0,31
C2	0,8	-0,07

Atual		
ID	x	y
C1	1,14	-0,22
C2	0,88	0,12

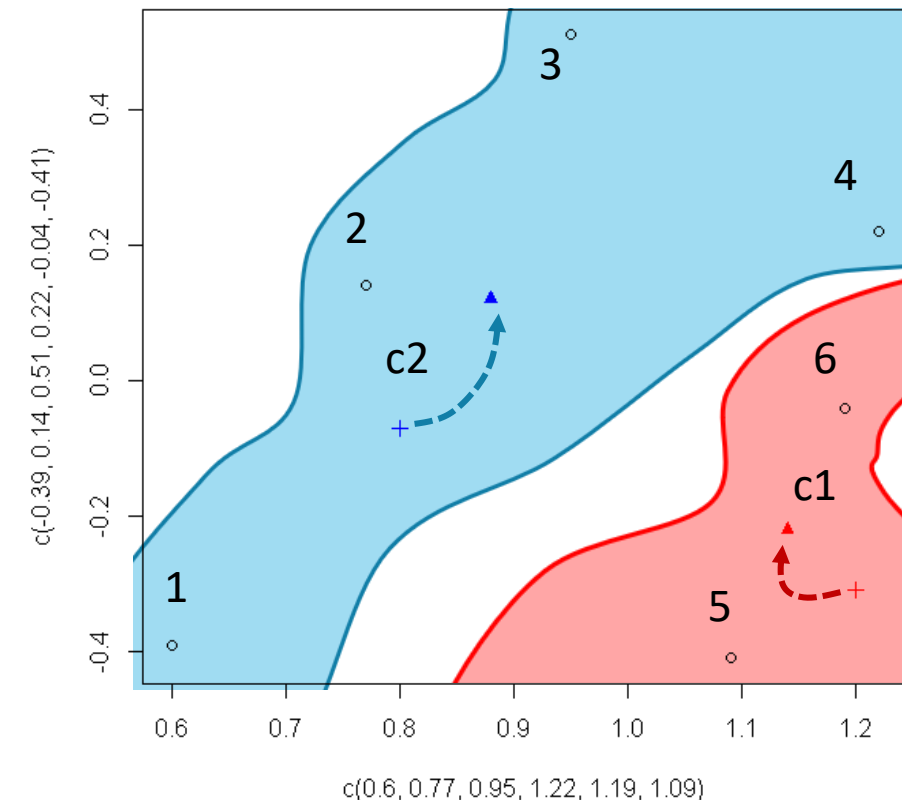
ID	Atributo A	Atributo B	C1	C2
1	0,6	-0,39	0,60	0,37
2	0,77	0,14	0,62	0,21
3	0,95	0,51	0,85	0,59
4	1,22	0,22	0,53	0,51
5	1,19	-0,04	0,27	0,39
6	1,09	-0,41	0,14	0,44



Recálculo da Distância

Atual		
ID	x	y
C1	1,14	-0,22
C2	0,88	0,12

ID	AtributoA	AtributoB	C1	C2	C1	C2
1	0,6	-0,39	0,60	0,37	0,56	0,58
2	0,77	0,14	0,62	0,21	0,51	0,11
3	0,95	0,51	0,85	0,59	0,75	0,39
4	1,22	0,22	0,53	0,51	0,44	0,35
5	1,19	-0,04	0,27	0,39	0,18	0,34
6	1,09	-0,41	0,14	0,44	0,19	0,57



Ajuste do Agrupamento

