

## Trabalho 3 – Agrupamento de dados

20443 – Tópicos em Informática X  
Mineração de Dados  
2018/2 - Turma A  
Prof. Dr. Murilo Naldi

[naldi@dc.ufscar.br](mailto:naldi@dc.ufscar.br)

# Divisão

- A grande empresa multinacional UNIBRAZIL decidiu que pretende dividir a gerencia da empresa em marcas menores e mais especializadas
- **Objetivo: levantar um perfil de clientes para cada uma das marcas**
- Vantagem: cada marca irá se especializar nos produtos que interessam ao seu grupo de clientes



## Pontos importantes

- Com novas marcas, as lojas serão especialistas em seus clientes, terão apenas nos produtos que interessam a eles, aumentando as vendas e reduzindo gastos com transporte e estoque
- Entretanto, dividir uma empresa em várias marcas pode ser um processo custoso
  - Gerenciar a marca, publicidade e funcionários também são especializados e por isso maior é o custo

# Objetivo do projeto

- Definir em quantas marcas a UNIBRAZIL deverá ser dividida e porque?
  - A partir da amostra da base de dados da empresa em “clientes.txt”
- Qual o percentual de clientes ficará sob a bandeira de cada marca?
- Qual o perfil de compra dos clientes de cada marca?
- Outras perguntas?





Prod1	Prod2	Prod3	Prod4	Prod5	Prod6	Prod7	Prod8	Prod9	Prod10
21	80	37	27	5	38	55	7	4	54
23	83	37	30	12	32	61	10	9	48
22	96	35	28	3	32	64	6	4	50
71	58	34	26	6	88	13	1	0	20
22	84	32	29	14	31	52	19	4	48
73	63	33	28	15	86	15	2	2	17
21	66	31	28	6	32	57	17	4	58
72	56	36	27	10	88	16	5	0	9
72	61	33	29	6	84	17	4	1	10
75	59	36	27	20	86	11	3	1	20
71	70	37	28	3	87	13	3	1	12
23	87	31	27	10	39	57	10	1	52
22	78	37	26	19	31	55	18	2	48
71	56	34	28	7	83	19	5	1	10
1	71	37	27	8	0	12	25	57	21
72	91	33	29	1	84	19	3	2	7
75	53	33	27	14	83	17	3	0	1
73	64	32	27	15	85	16	3	0	18
21	84	37	30	2	31	56	19	3	56
22	99	34	28	9	36	52	10	1	55
72	78	31	27	13	86	20	4	2	8
71	95	36	30	5	82	15	2	1	3
71	71	37	30	14	82	13	1	2	1
23	88	35	27	12	34	53	14	3	58
23	56	37	26	20	37	56	12	10	50
22	57	35	29	7	39	58	7	1	60
74	51	34	27	18	89	15	4	1	12

## Avaliação do trabalho

- Assim como a tarefa de agrupamento, o desenvolvimento e avaliação do trabalho serão subjetivos
- O grupo terá que convencer o CEO da UNIBRAZIL qual decisão é melhor para a empresa a partir da aplicação de técnicas de mineração de dados (pré-processamento, agrupamento e validação)
- Grupos de até 3 alunos

# Apresentação

- Objetivo:
  - Apresentar a aplicação e resultado das técnicas utilizadas em detalhes
  - Mostrar que o grupo domina o processo de descoberta de conhecimento
  - A nota de apresentação será dada individualmente

# Duração

- Cada grupo deverá apresentar o tema com *slides* entre 10-15 minutos (+perguntas)
- Apresentações que saiam deste intervalo terão suas notas penalizadas
- Os membros do grupo devem dividir o tempo de apresentação de maneira uniforme
- Avaliação individual de cada membro
- Desvios no horário serão penalizados





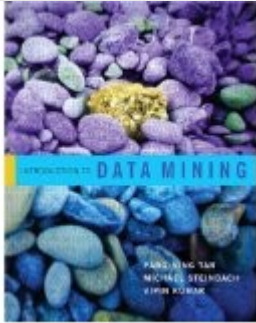
# Documentação

- O grupo deverá redigir um relatório (pdf) contendo o desenvolvimento do trabalho, passo a passo, como se ele fosse ser entregue ao CEO da empresa
- Também devem ser entregues os slides (em pdf)!!!!!!!
- Relatórios sem identificação de autores (nome e RA) terão nota 0 (zero)

# Notas

- A nota dada ao projeto será **individual** dividida da seguinte forma:
  - Domínio conteúdo
  - Didática
  - Bom uso do tempo
- O trabalho deverá ser entregue no AVA (apresentação + relatório em pdf) até o dia 12 de dezembro, 13:59 horas
- Apresentação será dia 12 de dezembro

# Referencias

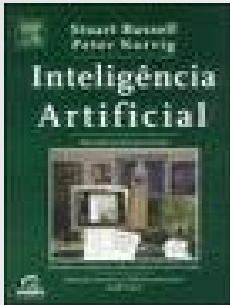


STEINBACH, M., KUMAR, V. TAN, P. Introdução ao Data Mining (Mineração de Dados). Edição 1. Ciência Moderna 2009. ISBN 9788573937619.

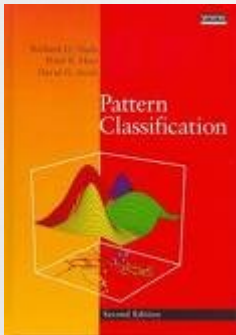


**Inteligência Artificial - Uma Abordagem de Aprendizado de Máquina. Katti Faceli, Ana Carolina Lorena, João Gama, André C. P. L. F. de Carvalho. Grupo Gen 2011**

# Referencias



**RUSSEL, S. and NORVIG, P.**  
**Inteligência Artificial: uma abordagem moderna. Editora Campus, 2003**



**Duda, R.O., Hart, P. E. and Stork, D. G. Pattern Classification (2nd Edition). Wiley-Interscience**

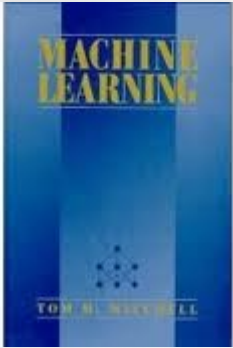


**SILVA, I.N., SPATTI, D.H. e**

**ELAUZING, B. A. Redes Neurais**



## Referencias



**MITCHELL, T. Machine Learning, McGraw Hill, 1997.**



**LINDEN, R. Algoritmos Genéticos. Brasport. 2008**