

Aplicação de mineração de dados ao problema de evasão de estudantes na UFC

Abelardo Vieira Mota

23 de maio de 2015

Resumo

é o abstract!!!!!!!

Sumário

1	Introdução	1
1.1	Contextualização	1
1.2	Objetivos	3
1.3	Organização do trabalho	3
2	Mineração de dados	3
2.1	Identificação de padrões	3
2.2	Predição	3
3	Evasão de estudantes	3

1 Introdução

1.1 Contextualização

O problema de evasão de estudantes refere-se ao abandono do curso pelo estudante antes de sua conclusão. De acordo com o Anuário estatístico da UFC de 2014, ano base 2013, o indicador "Taxa de Sucesso na Graduação"(TS) , definido como a proporção entre número de diplomados e número de ingressantes da graduação, esteve em 2013 com o menor valor desde que passou a ser monitorado. Já o indicador "Taxa de sucesso da graduação por curso" possui valor médio e mínimo para o curso . Esses indicadores apontam para que o problema da evasão de estudantes é uma realidade na UFC.

Os dados gerenciados pelos sistemas de informação da UFC podem conter informações que auxiliem o entendimento das causas desse problema e permitam que melhor sejam planejadas ações para solucioná-lo ou que estudantes com maior risco de evasão sejam identificados e recebam apoio da instituição. É nesse sentido, que a área de pesquisa denominada mineração de dados(data mining, em inglês), que estuda como extrair informações sobre dados, pode contribuir, fornecendo meios para a descoberta de informações relevantes a partir dos dados registrados pela UFC.

figura

figura

calcular

analisar

analisar

	Curso	Período	TS 2013
80	Ciências Sociais - Licenciatura	Noturno	6.8%
79	Redes de Computadores - Quixadá	Noturno	13.3%
78	Geografia - Bacharelado	Diurno	15.3%
77	Letras - Português-Alemão	Diurno	17.6%
76	Engenharia Metalúrgica	Diurno	18.3%
75	Ciências Econômicas - Sobral	Noturno	20.9%
74	Sistemas de Informação - Quixadá	Diurno	22.0%
73	Filosofia - Bacharelado	Noturno	24.3%
72	Matemática - Bacharelado	Diurno	24.4%
71	Engenharia Elétrica - Sobral	Diurno	25.0%
70	Letras - Português-Francês	Diurno	25.0%
69	Estatística	Diurno	25.0%
68	Física - Bacharelado	Diurno	25.0%
67	Educação Musical - Licenciatura	Diurno	25.0%
66	Design de Produto - Cariri	Noturno	25.5%
65	Engenharia da Computação - Sobral	Diurno	26.8%
64	Oceanografia	Diurno	27.8%
63	Física - Licenciatura	Noturno	31.3%
62	Geologia	Diurno	35.0%
61	Agronomia - Cariri	Diurno	36.5%
60	Engenharia de Teleinformática	Diurno	36.7%
59	Ciências Econômicas	Noturno	37.5%
58	Química - Licenciatura	Noturno	38.8%
57	Pedagogia	Noturno	40.8%
56	Economia Doméstica	Diurno	43.0%
55	Ciências Biológicas - Licenciatura	Diurno	43.5%
54	História - Bacharelado	Diurno	44.4%
53	Ciências Sociais - Bacharelado	Diurno	45.9%
52	Engenharia de Pesca	Diurno	50.0%
51	Engenharia Civil - Cariri	Diurno	51.9%
50	Química - Bacharelado	Diurno	52.0%
49	Engenharia Química	Diurno	52.1%
48	Computação	Diurno	53.2%
47	Filosofia - Licenciatura - Cariri	Noturno	53.7%
46	Design de Moda	Diurno	54.0%
45	Zootecnia	Diurno	54.5%
44	Engenharia de Produção Mecânica	Diurno	55.0%
43	Ciências Econômicas	Diurno	55.0%
42	Engenharia Mecânica	Diurno	55.9%
41	Engenharia Elétrica	Diurno	55.9%
40	Ciências Atuariais	Noturno	56.0%
39	Matemática - Licenciatura	Noturno	56.0%
38	Ciências Contábeis	Noturno	56.3%
37	Agronomia	Diurno	59.4%
36	Ciências Biológicas - Bacharelado	Diurno	59.6%
35	Ciências Sociais - Licenciatura	Diurno	60.0%
34	Administração	Diurno	62.0%
33	Biblioteconomia - Cariri	Diurno	63.3%
32	Arquitetura e Urbanismo	Diurno	64.1%
31	Psicologia - Sobral	Diurno	65.0%
30	Letras - Português-Italiano 2	Diurno	69.2%
29	Engenharia de Alimentos	Diurno	70.6%
28	Letras - Português-Espanhol	Diurno	71.1%
27	Letras - Português-Inglês	Diurno	71.4%
26	Administração	Noturno	72.8%
25	Publicidade e Propaganda	Diurno	74.5%
24	Biblioteconomia	Diurno	75.4%
23	História - Licenciatura	Diurno	78.6%

A grande tabela que ocupa mais de uma página!!!!

falar sobre os estudos sobre drop-out + data mining

1.2 Objetivos

O presente trabalho objetiva estudar o problema de evasão de estudantes na UFC a partir dos dados que seus sistemas de informação gerenciam, avaliando a aplicabilidade de técnicas de mineração de dados.

Para tanto é necessário que seja feita uma análise sobre a estrutura e qualidade dos dados disponíveis.

Também serão levantadas hipóteses sobre causas para o problema e será analisado se os dados as corroboram ou não.

1.3 Organização do trabalho

2 Mineração de dados

2.1 Identificação de padrões

2.2 Predição

[4] [6] [3]

sobre data mining - qual o super artigo básico?

3 Evasão de estudantes

[2] [5] [7] [1]

clustering, visualização

machine learning

Referências

- [1] RSJD Baker et al. Data mining for education. *International encyclopedia of education*, 7:112–118, 2010.
- [2] Gerben W Dekker, Mykola Pechenizkiy, and Jan M Vleeshouwers. Predicting students drop out: A case study. *International Working Group on Educational Data Mining*, 2009.
- [3] Pedro Domingos. A few useful things to know about machine learning. *Communications of the ACM*, 55(10):78–87, 2012.
- [4] Thomas M. Mitchell. *Machine Learning*. McGraw-Hill, Inc., New York, NY, USA, 1 edition, 1997.
- [5] Cristóbal Romero and Sebastián Ventura. Educational data mining: a review of the state of the art. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 40(6):601–618, 2010.

- [6] D. Sculley, Gary Holt, Daniel Golovin, Eugene Davydov, Todd Phillips, Dietmar Ebner, Vinay Chaudhary, and Michael Young. Machine learning: The high interest credit card of technical debt. In *SE4ML: Software Engineering for Machine Learning (NIPS 2014 Workshop)*, 2014.
- [7] Brandon Sherrill, William Eberle, and Doug Talbert. Analysis of student data for retention using data mining techniques. 2011.