

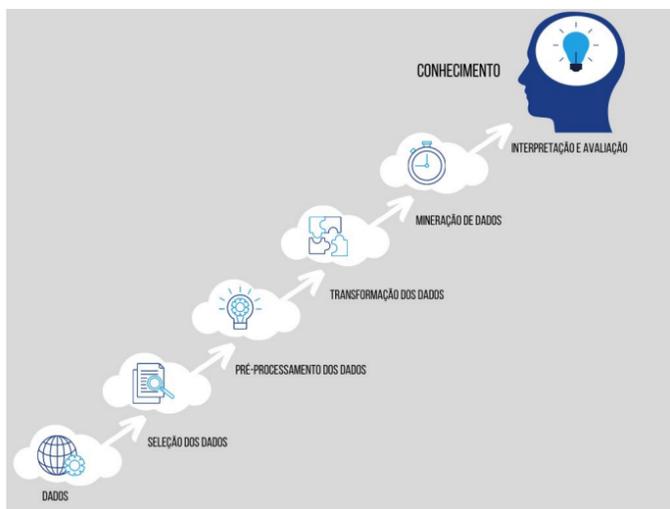
Mineração de dados educacionais para predizer o desempenho escolar dos estudantes de física ao longo do Ensino Médio.

Ricardo Dias Pacheco Martins¹, Carlos Diego Nascimento Damasceno², Seiji Isotani³

¹Pós-Graduando(a) em Computação Aplicada à Educação, ICMC-USP, <ricardomartins@usp.br>. ²Co Orientador, Universidade de São Paulo, <damascenodiego@usp.br>. ³Orientador, Instituto De Ciências Matemáticas e de Computação - USP, <sisotani@icmc.usp.br>.

INTRODUÇÃO

A começar do princípio que os dados estão sendo deixados de lado no campo educacional, o artigo pretende auxiliar os professores a decupar dados brutos através do conhecimento adquirido no campo da descoberta do conhecimento a partir de dados e aplicar o método do KDD (Knowledge Discovery Data) juntamente com o EDM (Educational Data Mining) para predizer de forma visual e prática como melhorar o desempenho escolar dos estudantes de Física do Ensino Médio ao longo do 1º, 2º e 3º anos.



Panorama dos passos que compõem o Processo de Descoberta de Conhecimento em Dados (KDD)

Fonte: Adaptação de FAYYAD et al. (1996).

OBJETIVOS

- Identificar possíveis recuperações ao longo das primeiras unidades e semestre.
- A outra indagação é a respeito de como prever caso o estudante esteja com chances de reprovação.
- Promover uma reflexão acerca da potencialidade do uso da mineração de dados no campo da educação

MATERIAIS E MÉTODOS

A partir dos dados obtidos, iniciou-se a etapa de desenvolvimento inspirado nos métodos KDD e EDM.



=== Stratified cross-validation ===
=== Summary ===

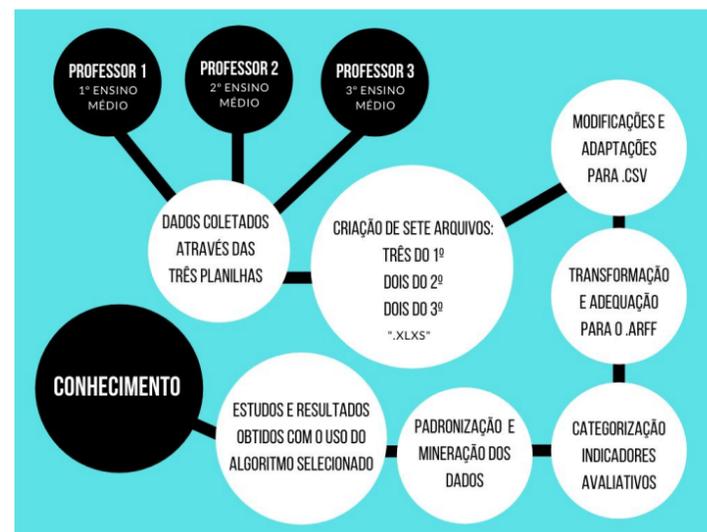
Correctly Classified Instances	225	90	%
Incorrectly Classified Instances	25	10	%
Kappa statistic	0.8557		
Mean absolute error	0.0448		
Root mean squared error	0.1964		
Relative absolute error	16.087	%	
Root relative squared error	52.7382	%	
Total Number of Instances	250		

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	217	87.1486	%
Incorrectly Classified Instances	32	12.8514	%
Kappa statistic	0.8225		
Mean absolute error	0.0622		
Root mean squared error	0.2147		
Relative absolute error	21.4297	%	
Root relative squared error	56.4076	%	
Total Number of Instances	249		
Ignored Class Unknown Instances	1		

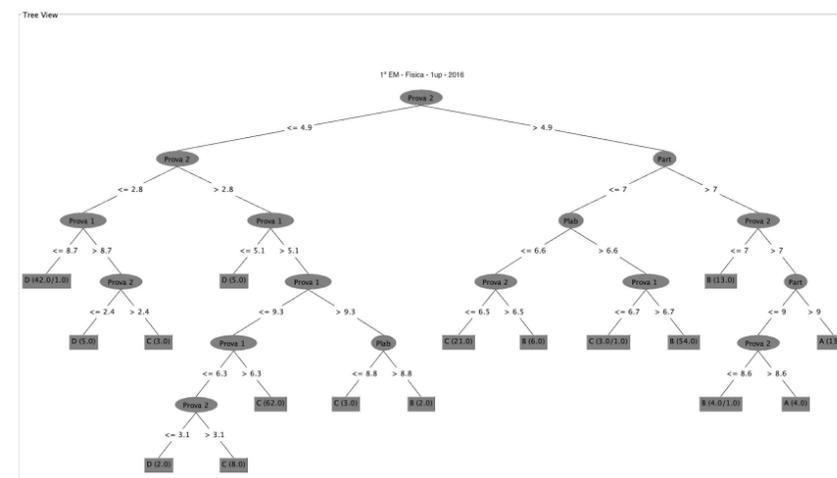
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	194	77.6	%
Incorrectly Classified Instances	56	22.4	%
Kappa statistic	0.6356		
Mean absolute error	0.0972		
Root mean squared error	0.2861		
Relative absolute error	39.0328	%	
Root relative squared error	81.3053	%	
Total Number of Instances	250		



Overview dos passos inspirados no processo do KDD.

Fonte: criado pelo autor



Árvore de decisão 1º EM - Física - 1ª unidade 2016

Fonte: gerado pelo software WEKA

CONSIDERAÇÕES FINAIS

- Considerando o algoritmo J48 para classificar os dados, os professores das séries do Ensino Médio poderão sugerir atenção para pontos importantes e que não estavam claros ao observar apenas o arquivo ".xlsx"
- Por fim, vale ressaltar que os inputs escolhidos fazem diferença para o sucesso da predição e interpretação dos resultados obtidos, assim como a escolha da árvore de decisão para fácil compreensão geral.
- Para trabalhos futuros vale debater qual o espaço da mineração de dados dentro da escola e voltar o questionamento de como os dados podem influenciar no poder de decisões futuras em diversos aspectos pedagógicos.

REFERÊNCIAS

Baker, R.S.J.d. (2010). Data Mining for Education. McGraw, B., Peterson, P., Baker, E. (Eds.) International Encyclopedia of Education (3rd edition). Oxford, UK: Elsevier.

Baker, R.S.J.d., I. S. d. C. A. (2011). Mineração de dados educacionais: Jornada de Atualização em Informática na Educação - JAIE 2012 25/29 Oportunidades para o Brasil. Revista Brasileira de Informática na Educação, 19(2).

Fayyad, Usama; Piatetsky-Shapiro, Gregory; Smyth, Padhraic. (1996). From Data Mining to Knowledge Discovery in Databases. Aaai Press, p.37-54.

Witten, I. H., Frank, E. (2005). "Data Mining: Practical Machine Learning Tools and Techniques". 2nd edition, Morgan Kaufmann Publishers, San Francisco, CA.