

Integração de um mecanismo de Mineração de Dados Educacionais ao Moodle

Leandro C. Santana¹, Alexandre M. Maciel¹, Rodrigo L. Rodrigues²

¹Universidade de Pernambuco (UPE)
Programa de Pós-Graduação em Engenharia da Computação (PPGEC)

²Universidade Federal Rural de Pernambuco (UFRPE)
Departamento de Educação (DEd)

{lsc, amam}@ecomp.poli.br, rlr@ded.ufrpe.br

Abstract. *Virtual Learning Environments (VLE) are an example of new systems being developed as part of educational environments. Moodle is an example of a VLE distribution which offers courses and training to several persons. In 2012 the Moodle plugin "Avatar Educação" was developed and with it is possible to create personalized notifications and messages within Moodle. This paper proposes an integration mechanism for the "Avatar Educação", based on decision trees, which offers a better control and distribution of data within the plugin. This way, we provide a better interactivity between the users and the VLE.*

Resumo. *Um ambiente educacional a cada dia vem incorporando software em suas atividades, os Ambientes Virtuais de Aprendizagem (AVA) são exemplos desse software. O AVA, é uma forma de oferecer cursos e treinamentos e tem como uma das suas distribuições, o Moodle. Em 2012 surgiu o "Avatar Educação", um plugin desenvolvido para notificações e mensagens personalizadas integrada ao Moodle. Este trabalho objetiva a integração de um mecanismo baseado em árvore de decisão integrado ao Avatar Educação afim de oferecer respostas inteligentes a comunidade de Ensino a Distância. Desta forma é possível obter uma interatividade melhor entre os usuários e o ambiente.*

1. Introdução

O processo de ensino, seja presencial ou ensino a distância (EAD), leva em consideração a interação entre os atores envolvidos no processo: alunos, professores e ambiente. A interação entre esses atores se torna importante dentro do processo, pois proporcionam maior entendimento na relação entre professores e alunos, o acompanhamento é de forma cooperativa e a rede de ensino é influenciada positivamente (SILVA, 2001).

Na pesquisa de Maciel (2014), intitulada, Avatar Educação, deu início a um processo de melhoria de interação entre os atores envolvidos no processo educacional, porém apresentou limitações de funcionalidades. As consultas realizadas pelo Avatar ao banco de dados foram desenvolvidas de forma estática e não agregam nenhum

mecanismo inteligente na busca por informação. Isto acarreta um desinteresse gradativo dos alunos e em longo prazo deixam de contribuir para o incremento na interatividade do ambiente (Romero et al., 2008).

Diante desta problemática, este trabalho tem como objetivo apresentar a integração de um mecanismo de mineração de dados educacionais ao ambiente educacional Moodle, a fim de melhorar a interação com os usuários do ambiente.

2. Mineração de Dados Educacionais

Romero, Ventura e García (2008) afirmam que na EAD a quantidade de armazenamento de dados vem crescendo, e isso se dá ao fato que as interações como acessos a fóruns, perguntas e respostas na área de questionários e comunicação entre participantes. Diante dessas atividades um grande volume de dados é desenvolvido dificultando uma análise de forma manual.

As técnicas de mineração de dados foram desenvolvidas a muitos anos para ser aplicado em áreas de marketing, vendas e apoio financeiros (Fayyad et al., 1996; Witten et al., 2011) mais recentemente, vem sendo aplicada na área educacional (Baker, 2011). Na área de educação, Zorrilla et al. (2005) apontam que estas técnicas podem ser aplicadas para diferentes atores que são: professores, estudantes e os gestores. Onde cada um tem sua característica e possui seu ponto de vista.

Diversas abordagens e algoritmos são utilizados na mineração de dados: regressão linear, redes neurais, máquinas de vetores de suporte e árvores de decisão. Romero et al. (2008) salientam que dentre os diversos métodos de mineração, destacam-se as árvores de decisão e as redes neurais, onde as árvores de decisão abordam um conjunto de dados e uma classificação seguindo um caminho que satisfaça as condições, onde inicia-se pela raiz até a folha da árvore. Abordaremos mais especificamente as técnicas de árvore de decisão foco deste trabalho, tendo em vista que a técnica tem melhor resultado com base em experimentos em base de dados educacionais (GOTTARDO, 2014).

2.1. Árvore de Decisão

Árvore de decisão é uma técnica de aprendizado de máquina que aborda uma estrutura de árvore para tomada de decisão. A árvore possui vários nós, onde cada nó corresponde a uma tomada de decisão baseada por um determinado dado de entrada. A ligação entre um nó e outro é chamado de ramo, e o último nó no qual não tem mais ligação é chamado de folha, que é o retorno da árvore (QUINLAN, 1993).

A Figura 1 ilustra a estrutura de uma árvore de decisão onde cada nó é realizado uma decisão, e ao chegar na folha, é retornada uma resposta para a mineração. No caso proposto deste trabalho, a resposta da árvore são as opções da classe de entrada no classificador.

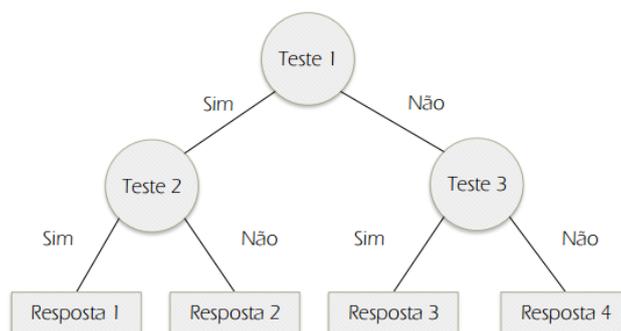


Figura 1. Estrutura de uma árvore de decisão (Rolim, 2014).

Árvore de decisão tem sido largamente utilizada nos trabalhos de mineração de dados educacionais (WANG, 2009). Por este motivo exploramos dos dois principais algoritmos que envolvem os estudos com árvore de decisão, são eles C4.5 (J48) e Random Forest.

3. Experimentos e Resultados

Os experimentos deste trabalho foram desenvolvidos em dois cenários: o primeiro foi uma experiência com 79 estudantes e com 6 atributos de uma base de dados do Moodle e foi relatada em Cavalcanti *et al.* (2014). O experimento do segundo cenário foi realizado com uma base de dados ampla e vários cursos e disciplinas com 1254 alunos e 19 atributos. A característica dos atributos pode ser vista na Teoria da Interação proposta por Moore (1989), onde o mesmo define os dados em dimensões. Neste trabalho as dimensões são chamadas da seguinte forma: Perfil de Uso do AVA, Interação Estudante – Estudante e Interação Bidirecional. A relação Dimensão x Atributos é dada a seguir:

- Perfil de Uso do AVA
 - ✓ Número total de acesso ao AVA;
 - ✓ Número de postagens de outros participantes lidas em fóruns;
 - ✓ Número total de revisões em postagens anteriores realizadas em fóruns;
 - ✓ Número de sessões de chat que o estudante participou;
 - ✓ Número de mensagens enviadas ao chat;
 - ✓ Número de questões respondidas;
 - ✓ Número de questões respondidas corretamente;
 - ✓ Tempo médio decorrido entre os diversos acessos ao sistema;
 - ✓ Número de dias transcorridos entre o início do curso e o primeiro acesso do estudante no AVA.
- Interação Estudante - Estudante
 - ✓ Número total de respostas postadas em fóruns referindo-se a postagens de outros estudantes;

- ✓ Número de postagens em fóruns de outros estudantes que fazem referência a postagem do estudante;
 - ✓ Número de mensagens recebidas de outros estudantes durante a realização do curso;
 - ✓ Número de mensagens enviadas a outros estudantes durante a realização do curso.
- Interação Bidirecional
 - ✓ Número de postagens do estudante que tiveram respostas feitas por professores ou tutores do curso;
 - ✓ Número de postagens de professores ou tutores que tiveram respostas feitas pelo estudante;
 - ✓ Número de mensagens enviadas ao professor/tutor durante a realização do curso;
 - ✓ Número de mensagens recebidas do professor/tutor durante a realização do curso.

O objetivo da previsão é o resultado final obtido pelo estudante no curso. Representa a classe objetivo de classificação. Nesse experimento foi utilizado técnicas de mineração Randon Forest e J48 por meio da ferramenta Weka. Na Tabela 1 é possível visualizar a taxa de acerto usando as técnicas.

RF	J48
94,37%	94,69%

Tabela 1. Acurácia dos Algoritmos

Diante dos resultados, chegamos a uma performance de 94,69% do algoritmo J48. Para ilustrar melhor o resultado, a Figura 2 e 3 mostram os resultados e matriz de confusão do experimento.

Correctly Classified Instances	1195	94.691 %
Incorrectly Classified Instances	67	5.309 %
Kappa statistic	0.6927	
Mean absolute error	0.0664	
Root mean squared error	0.2195	
Relative absolute error	37.6181 %	
Root relative squared error	74.0038 %	
Total Number of Instances	1262	

Figura 2. Resultado do Algoritmo J48

=== Confusion Matrix ===

```

a    b  <-- classified as
87   36 |    a = AP
31 1108 |    b = REP

```

Figura 3. Representa o resultado da Matriz de Confusão j48

Para todos os experimentos dos cenários se utilizou da normalização padrão, min-máx e z-score.

3.1. Integração ao Moodle

Após realização dos experimentos e escolha da melhor técnica de mineração em relação a série de dados escolhidas, a integração ao Moodle aconteceu da seguinte maneira:

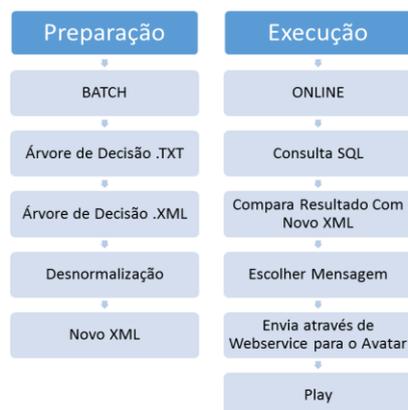


Figura 4. Fases de desenvolvimento do mecanismo de inteligência

A Figura 4 caracteriza a etapa de desenvolvimento em dois cenários distintos, onde inicia-se com a preparação dos dados, em seguida a fase de execução. A preparação dos dados foi executada localmente (localhost), onde foi criado o arquivo .TXT contendo a árvore de decisão gerado pela saída do Weka. O arquivo .TXT foi importado ao plugin WekaTextToXML e gerado um .XML logo em seguida. Após o arquivo XML ter sido gerado, o próximo passo foi de desnormalizar os dados, tendo em vista que os dados foram normalizados inicialmente para poder ser aplicado as técnicas de mineração. Após essa desnormalização, um novo arquivo XML foi gerado.

O cenário da execução dos dados se dá por meio do servidor online do Moodle, onde o aluno faz o *login* com seu ID e o Avatar Educação faz consultas na base de dados do aluno, fazendo comparações ao novo XML (árvore de decisão) escolhendo a mensagem que caracteriza o perfil do aluno e executando o áudio da mensagem por meio da voz sintetizada.

As mensagens foram divididas em duas categorias, mensagens motivadoras e mensagens de alerta como pode ser vista na Tabela 2 e na Tabela 3. Ambas transmitem estímulos para o estudante seguir motivado no curso. As mensagens motivadoras, são selecionadas quando a folha da árvore tem como resultado “aprovado”, já as mensagens de alerta, tem uma propriedade de motivar o aluno a melhorar dentro da disciplina ou curso, e estas mensagens são selecionadas quando a folha da árvore tem como resultado “reprovado”.

MENSAGEM MOTIVADORA	
NR_MSG_REC_PROF	Parabéns pela comunicação com seu tutor! Continue assim!
NR_QUESTOES_RESP	Continue fazendo os exercícios! Isso é ótimo para ajudar no seu bom desempenho!
NR_MSG_CHAT	Parabéns pelo uso constante do chat. Continue assim! As trocas com tutores e colegas podem ser muito

	enriquecedoras.
RESP_PROF_EST	Parabéns por sua interação no fórum! Continue comunicando-se com seu professor, tutor e colegas!
TEMPO_DECORRIDO	Você vem acessando o ambiente regularmente. Isso vai ajudar no seu rendimento!
TEMPO_DECOR_PRIM_LOGIN	Sua frequência de acesso no ambiente está muito boa! Continue sempre assim!
NR_POSTS_REC	Continue sempre acessando os fóruns, compartilhando experiências com outros alunos!
NR_TOT_POST	Continue sempre interagindo no fórum! Sua participação está ótima!
NR_QUESTOES_ACERT	Continue respondendo sempre as atividades! Ótimo trabalho!
NR_MSG_ENV_EST_PROF	Percebo que você sempre busca seu tutor para tirar as dúvidas! Isso é ótimo! Continue assim!
NR_MSG_REC_EST	Procure sempre ajudar seus colegas no que precisarem.
NR_LOGINS	Continue sempre acessando o ambiente. Muito bom isso.
NR_CHATS	Você está tendo uma ótima comunicação com os participantes, muito bem.

Tabela 2. Mensagens motivadoras geradas pelo Avatar

MENSAGEM DE ALERTA	
RESP_EST_PROF	Procure verificar sempre os fóruns. Eles são muito importantes! Bons estudos!
NR_TOT_POSTS	Você já interagiu no fórum hoje? Verifique as novidades. Talvez seu amigo tenha alguma dúvida que você possa ajudar!
NR_MSG_ENV_EST	Lembre-se que os chats são semanais. Aproveite e tire as dúvidas, se houver!
NR_TOT_REV	A interação com o ambiente é muito importante! Nunca deixe de acessá-lo!
NR_QUESTOES_ACERT	Procure reler as anotações de aulas, assistir aos vídeos e seguir as orientações de estudo do professor. Depois, tente refazer os exercícios.
NR_MSG_REC_EST	Procure ajudar os demais colegas. Verifique se eles precisam de sua ajuda nos fóruns!
NR_QUESTOES_RESP	Não deixe para a última hora. Siga as orientações do professor e responda o quanto antes os exercícios!! Bons estudos!
NR_MSG_REC_PROF	Você se comunicou com seu tutor recentemente? Verifique se ele tem novidades para você!
TEMPO_DECOR_PRIM_LOGIN	A regularidade de sua participação é essencial para o sucesso do curso, nunca deixe de acessar a plataforma!

TEMPO_DECORRIDO	Sempre procure acessar o ambiente, isso ajudará seu desempenho nas atividades!
NR_LOGINS	Acesse o ambiente regularmente, assim vai melhorar seu desempenho.
NR_CHATS	Converse com seus amigos, veja se eles precisam de sua ajuda.

Tabela 3. Mensagens de alerta geradas pelo Avatar

3.1.1. Moodle/Avatar

Essa interação dar-se por meio de um personagem animado com voz sintetizada, o Avatar Educação, que pode ser visualizado na Figura 5.

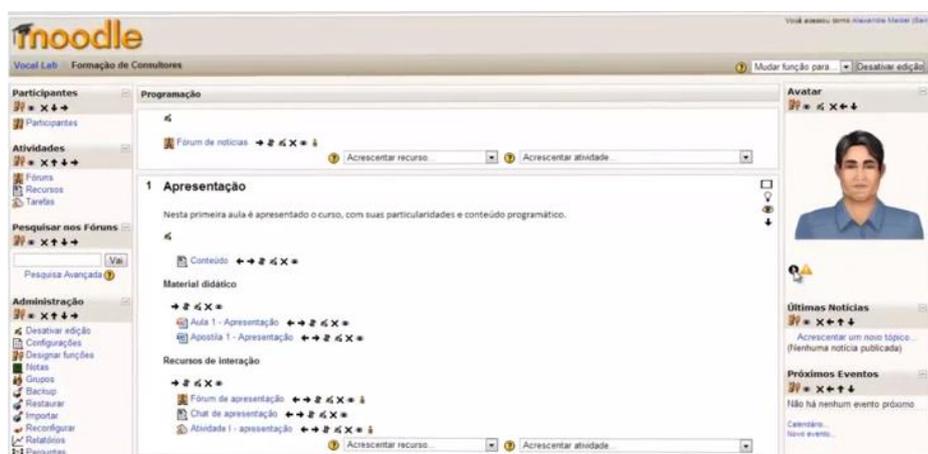


Figura 5. Tela do Ambiente Moodle com o Avatar Educação (Maciel, 2014).

O uso das técnicas de mineração por meio da árvore de decisão favorece o Avatar Educação ao adicionar uma camada mais robusta de processamento de informações subjetivas e contextuais. Na versão anterior, o levantamento de escolha das mensagens a ser enviadas ao usuário se dava baseado em escolhas pré-definidas e que se mostravam não-flexíveis, causando um efeito de mecanização na comunicação. Com a adição da árvore de decisão, conseguimos introduzir flexibilidade e customização individual na comunicação com o usuário, o que tende a aumentar o grau de engajamento do usuário com o sistema.

Essa etapa é a manipulação das classes do plugin Avatar Educação e incluir o mecanismo inteligente em sua estrutura. A estrutura do Avatar Educação antes da inserção do mecanismo de inteligência pode ser vista na Figura 6.

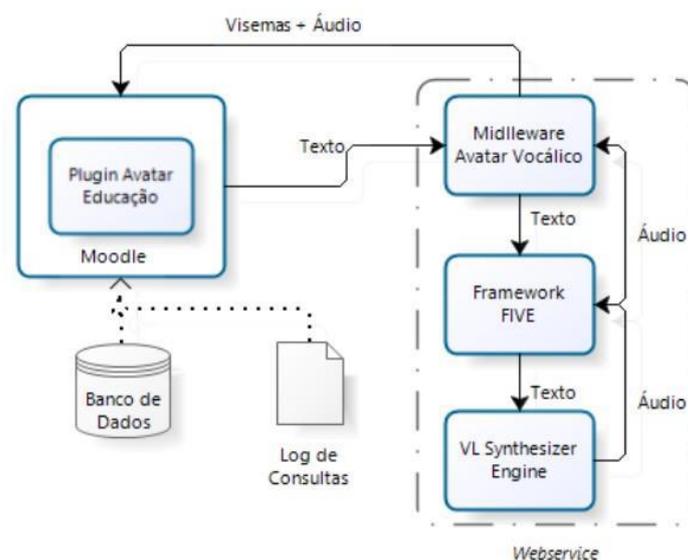


Figura 6. Arquitetura do Avatar sem o mecanismo inteligente

Depois do mecanismo de inteligência ser inserido ao Avatar Educação, a estrutura recebeu uma atualização, que pode ser vista na Figura 7. A mensagem inteligente está diretamente ligada ao mecanismo de mineração, recurso que será inserido ao Avatar Educação para realizar a sintetização da mensagem e reprodução do áudio.

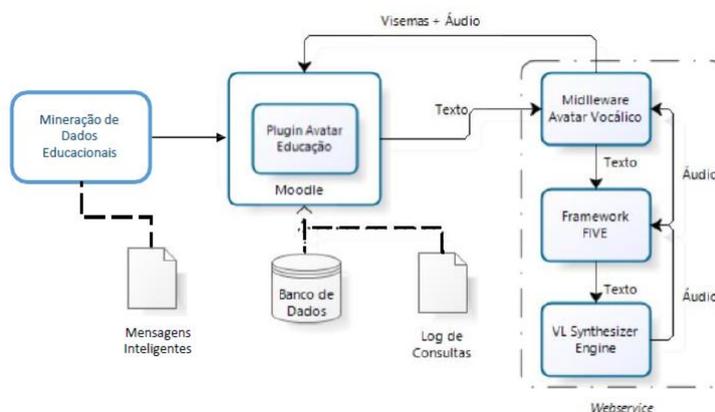


Figura 7. Arquitetura do Avatar com o mecanismo inteligente

Para ilustrar a etapa de desenvolvimento, a Figura 8, mostra as etapas até a execução do mecanismo.

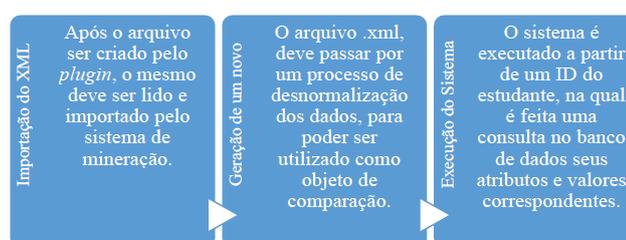


Figura 8. Etapas de execução do desenvolvimento do mecanismo

A Execução do Sistema, inicia-se através do aluno fazendo o *login* no AVA, ao fazer este procedimento, o Avatar Educação verifica se nas configurações a opção de

mensagens inteligentes está marcada (a configuração pode ser vista na Figura 9). Deste modo, o sistema verifica o ID do aluno e faz comparações dos atributos do aluno logado com os atributos da árvore de decisão, percorrendo a árvore verificando se o aluno foi aprovado ou reprovado.

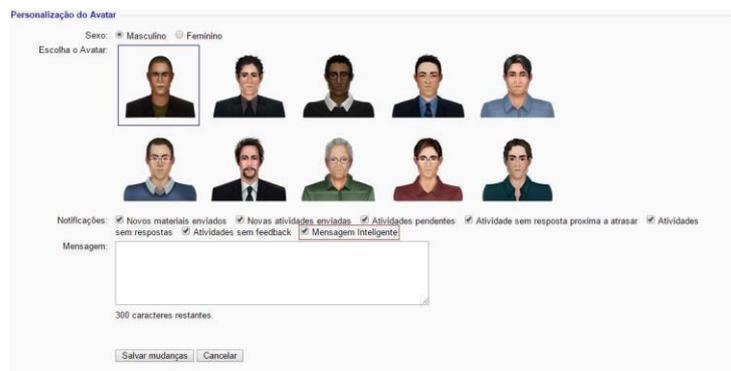


Figura 9. Tela de configuração do Avatar Educação

É importante destacar que o mecanismo de inteligência está voltado para o estudante, ou seja, as mensagens inteligentes e manipulação dos dados é em benefício ao aluno do AVA.

4. Conclusão

Este trabalho apresentou um modelo de aplicação na área de mineração de dados educacionais, por meio de um mecanismo ao plugin Avatar Educação. Foi realizado em um ambiente Moodle, um ambiente facilitador de comunicação durante o processo pedagógico, pois oferece uma contribuição para a comunicação entre os professores/tutores e alunos de uma disciplina ou curso.

Analisando-se os melhores resultados dos cenários estudados ao longo dessa pesquisa, percebe-se que as taxas de acerto das técnicas variaram entre 80% e 94%, podendo ser consideradas expressivas no contexto educacional. Afirmação que se baseia no resultado de Hämäläinen e Vinni (2011) que, por meio de pesquisas relacionadas a desempenho obteve uma taxa média de acerto de 72% entre as pesquisas.

Do ponto de vista tecnológico é uma nova funcionalidade integrada ao Avatar que torna as mensagens mais dinâmicas, agregando um valor a ferramenta educacional, visando a busca da interatividade maior com os alunos, fazendo com que os mesmos se dediquem em suas atividades no ambiente.

Referências

- 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).
- Cavalcanti, L.; Maciel, A.; Rodrigues, R. (2014). Avaliação do Perfil de Uso no Ambiente Moodle Utilizando Técnicas de Mineração de Dados. Simpósio Brasileiro de Educação a Distância, Dourados, 2014.

- Fayyad, U.; Piatetsky-Shapiro, G.; Smyth, P. (1996). Knowledge Discovery and Data Mining: towards a unifying framework. In: Second International Conference ON KD & DM.
- Gottardo, E., Kaestner, C. A. A., Noronha, R. V. (2014). Estimativa de Desempenho Acadêmico de Estudantes: Análise da Aplicação de Técnicas de Mineração de Dados em Cursos a Distância. RBIE, Vol 22, No 1, p. 45-54.
- Hämäläinen, W., Vinni, M. (2011). Classifiers for Educational Data Mining. In: Romero et al. Handbook of Educational Data Mining. Flórida, CRC Press, p. 57-71.
- Maciel, A. M. A.; Rodrigues, R. L.; Carvalho, E. C. B. (2014). Desenvolvimento de um Assistente Virtual Integrado ao Moodle para Suporte a Aprendizagem Online. Simpósio Brasileiro de Educação a Distância, Dourados.
- Moore M. G. (1989). Three Types of Interaction. The American Journal of Distance Education, 3(2):1-6.
- Quinlan, J. R. (1993). C4.5: Programs for machine learning. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
- Romero, C., Ventura, S. e García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. Computers and Education, 51 (1), pp. 368-384.
- Silva, E., Menezes, E. (2001). Metodologia da Pesquisa e Elaboração de Dissertação. Florianópolis: Universidade Federal de Santa Catarina – UFSC.
- Wang, Y. Tseng, M. Liao, H. (2009). Data mining for adaptive learning sequence in English language instruction. Expert Syst. Appl .J.,vol.36, pp. 7681-7686.
- Witten, I.H., Frank, E., Hall, M.A. (2011). Data Mining: Practical Machine Learning Tools and Techniques. San Francisco: Morgan Kaufmann, 3 ed.
- Zorrilla, M. E., Menasalvas, E., Marin, D., Mora, E., Segovia, J. (2005). Web usage mining project for improving web-based learning sites. In Web Mining Workshop. Cataluna (pp. 1-22).. 449-460.