

Curso: Sistemas de Informação

Equipe:

Professor Coordenador e Orientador: Adriano Araújo Santos/MSc

Alunos: André Iago Maracajá Albuquerque
Breno Lacerda de Alustau Paiva
Inácio Fabrício Nunes de Sousa
Michael Clementino de Almeida

ANÁLISE E DESENVOLVIMENTO DE MÉTODOS PARA MINERAÇÃO DE DADOS EDUCACIONAIS

Relatório de Pesquisa

**Campina Grande-PB
2013**

ADRIANO ARAÚJO SANTOS

**ANÁLISE E DESENVOLVIMENTO DE MÉTODOS PARA
MINERAÇÃO DE DADOS EDUCACIONAIS**

Relatório de Pesquisa apresentado ao Núcleo de Pesquisa e de Extensão (Nupex) do Centro de Ensino Superior e Desenvolvimento (Cesed) de acordo com o que preconiza o regulamento.

Campina Grande-PB
2013

Sumário

Capitulo 1 Introdução.....	1
Capitulo 2 Fundamentação Teórica	7
Capitulo 3 Trabalhos Correlatos	113
Capitulo 4 Metodologia	16
Capitulo 5 Desenvolvimento.....	18
Conclusão.....	22
Referências	23

Lista de Figuras

Figura 1: Disponibilização dos Arquivos do Módulo de Agente Inteligente Educacional -----	8
Figura 2: Tela inicial do MySQL Workbench -----	20
Figura 3: Módulo de Cálculo de Similaridade -----	21

Lista de Tabelas

Tabela 1: Descrição do módulo ----- 19

Capítulo 1

Introdução

Descrição do tema

A necessidade da expansão da educação e a incorporação de tecnologias da informação e comunicação, nos diferentes níveis de educação, fazem com que seja necessária investigação e adoção de novas formas de ensino-aprendizagem (SANTOS et al, 2012). Assim sendo, conforme JAIED (2012), a Educação a Distância (EAD) vem ganhando forças, principalmente porque se trata de uma modalidade de educação distinta dos modelos tradicionais e, quiçá, presenciais.

O forte crescimento do mercado da EAD, sua aceitação e disseminação têm tomado proporções exponenciais e alcançado diversos públicos, que navegam entre ensino fundamental, médio, técnico, superior e, até mesmo, pós-graduações e, em consequência a isso, cada vez mais pessoas passam a participar desse mercado, nas mais variadas funções: professores, alunos, produtores de conteúdo, pedagogos etc. e, muitas vezes, desempenhando diferentes papéis simultaneamente (FORMIGA, 2009).

Mesmo sabendo da velocidade e proporção com as quais esse mercado vem se expandindo, os recursos humanos não vêm acompanhando esse crescimento com a mesma agilidade. Segundo Maia (2008), ainda não existe uma total compreensão sobre os novos papéis que cada participante (professores, alunos, autores etc.) desempenha seus direitos, e suas responsabilidades em todo o novo processo de ensino-aprendizado.

Maia (2008) complementa seu pensamento afirmando que a) professores, tutores, autores acabam se confundindo com a sua nova função pedagógica; b) alunos desconhecem seu novo papel, pois não foram acostumados a ser um participante ativo no que diz respeito ao autogerenciamento do aprendizado; c) pedagogos não acompanham o desenvolvimento tecnológico e, por fim, d) tecnólogos abominam a pedagogia, andragogia e heutagogia, das quais sequer ouviram falar.

A cada dia se torna mais evidente a necessidade de profissionalizar aqueles que trabalham com educação no uso das novas tecnologias e, construir materiais que acompanhem as novas necessidades do processo de ensino-aprendizagem.

Nos últimos anos, muitos investimentos têm sido feito por instituições de ensino neste sentido e, segundo Filatro (2008) cursos de Design Instrucional surgiram para suprir a lacuna nessa formação, mas, no Brasil, comparado aos índices internacionais, há muito que evoluir.

Comprovando esse fato, a EAD é utilizada no âmbito do ensino fundamental e médio em países de língua inglesa, tais como EUA, Reino Unido, Canadá, Austrália. Na Nova Zelândia, é forte o movimento de home schooling,

home study ou home education. Nos EUA, é crescente a oferta de cursos a distância on-line para os alunos da high school.

Segundo Maia (2008), no Brasil, a EAD pode ser utilizada com restrições no ensino fundamental e médio, para crianças e adolescentes, e na educação adulta para esses níveis. No entanto, existe uma concentração de EAD para o ensino superior, desde que esta seja autorizada pelo MEC.

Na esfera privada, diversas empresas tais como IBM, Sun, Toyota etc., criaram universidades corporativas. JAIED (2012) classifica essas universidades como uma ferramenta estratégica projetada para assistir sua organização-mãe a atingir sua missão, conduzindo a atividades que cultivem aprendizado individual e organizacional, conhecimento e sabedoria. Além disso, essas universidades diferem do clássico departamento de treinamento, pois oferecem cursos com certificado acadêmico.

Existem inúmeras vantagens na utilização da EAD pelas organizações. Dentre as vantagens existentes na adoção dessa modalidade está a economia financeira (tanto para funcionários quanto para as organizações) porque, além da infraestrutura da própria empresa ser aproveitada, a quantidade de alunos e o perfil desses podem ser previstos com exatidão, o que facilita o planejamento do curso.

Vendo que este mercado é bastante atrativo, empresas de todo o mundo estão se especializando na criação de materiais que compõem a educação a distância, se classificando como fornecedores de EAD. Segundo Maia (2008), esses fornecedores de serviços e produtos de EAD produzem e vendem software, material didático, conteúdo e cursos completos e ainda prestam consultoria, oferecem hospedagem e serviço de logística, equipamentos e todo tipo de material que possa ser utilizado no processo de EAD.

No Brasil, em 27 de maio de 1996, foi oficialmente criada a Secretaria de Educação a Distância (SEED) e, entre as ações desta Secretaria estão a criação do canal TV Escola, apresentação do documento-base do programa “Informática na Educação” e várias outras iniciativas de educação on-line, que têm como objetivo promover o uso pedagógico da informática na rede pública de educação básica (MEC, 2012).

Além das iniciativas citadas, o Governo Federal do Brasil tem criado e apoiado diversas iniciativas e projetos de EAD, tais como: Domínio Público – Biblioteca Virtual, E-ProInfo, E-Tec Brasil, Sistema de Universidade Aberta do Brasil. Em algumas de suas instituições têm sido adotados projetos de educação que utilizam EAD. Exemplos dessas instituições são: o Tribunal de Contas da União, a Secretaria da Fazenda do Estado de Pernambuco, a Marinha e o Serviço Federal de Processamento de Dados (SERPRO) (MEC, 2012).

Na educação on-line, segundo Roesler (2011), os alunos vivenciam um contexto cibercultural, a partir do qual preponderam as relações sociais. Assim como a EAD, as Redes Sociais on-line e sites para construção de conhecimento colaborativo têm tomado uma grande proporção no dia a dia das pessoas, tornando as informações cada vez mais acessíveis formando uma Inteligência Coletiva (LÉVY, 1998). O problema nesse contexto é identificar quais as

informações e as pessoas que participam da construção desse conhecimento e que são relevantes para um interesse específico.

Acoplando Inteligência Coletiva aos sistemas de aprendizagem, utilizando, por exemplo, o Moodle (MOODLE, 2011), que permite o armazenamento de materiais, como também a comunicação entre alunos e professores a partir de chats e fóruns de discussão. Com o foco no apoio ao ensino presencial, acredita-se que fomentará a curiosidade dos alunos e dos professores a procurar materiais de outras disciplinas, ou até mesmo alunos e professores de semestres anteriores, a fim de comentarem assuntos de interesse de ambos.

Em Carvalho (2012), é proposto o conceito de Inteligência Coletiva Inteligente, que, em linhas gerais, permite o acesso às informações e às pessoas envolvidas no processo de construção do conhecimento de forma inteligente, por meio de recomendação de materiais e pessoas que colaborem para o enriquecimento individual.

O conceito surgiu a partir do momento em que foi observado que o grande desafio na mediação de grupos de aprendizagem em educação está na sobrecarga de trabalho do instrutor no acompanhamento do grande número de mensagens trocadas nesses ambientes, acumulando-se às competências tradicionais desse profissional (Carvalho, 2012).

Existem diversas formas de interação em ambientes virtuais de aprendizagem e, segundo Baker, Isotani, e De Carvalho (2011), apesar de muitos ambientes possuírem ferramentas de chat, fóruns de discussão e outras formas de interação, são poucos os trabalhos científicos que analisam os dados obtidos por essas ferramentas e correlacionam o conteúdo das conversas com a aprendizagem dos alunos.

Nesse contexto, surge uma área de atuação, relativamente, nova no Brasil chamada Mineração de Dados Educacionais (EDM) (Baker, Isotani, e De Carvalho, 2011), tendo como objetivo o desenvolvimento de métodos para explorar conjuntos de dados coletados em ambientes educacionais.

A EDM se utiliza da expertise da Mineração de Dados para analisar as relações existentes entre a abordagem pedagógica e o aprendizado do aluno, utilizando diversos algoritmos e ferramentas da Inteligência Artificial, considerando as diversas variáveis que influenciam a aprendizagem, a exemplo das variáveis relacionadas ao comportamento do estudante (Pavlik et al., 2008).

Segundo Lévy (2012), hoje, mesmo com as escolas de ensino à distância e outras formas de ensino, o destino final é o aprendizado social totalmente integrado ao ambiente da sociedade. É preciso, sobretudo, pensar em uma escola social.

Delimitação do tema

Com base nas informações expostas, é notória a necessidade de criação de estratégias que acompanhem a evolução educacional e que auxiliem o processo de ensino-aprendizagem, com o intuito de investigar o uso de técnicas de

Inteligência Artificial e como essas técnicas podem contribuir para esse propósito.

Um dos fatores observado foi a crescente adoção e aceitação das redes sociais no contexto atual da Cibercultura (LÉVY, 2008) e as possibilidades existentes do uso de Inteligências Coletivas (LÉVY, 2012) aplicadas a sistemas de aprendizagem, com o uso de sistemas de gerenciamento de aprendizagem e recomendação.

Identificação do foco da pesquisa fazendo constar a definição do problema

A pretensão inicial do trabalho consiste em desenvolver um agente inteligente que possibilite recomendações entre os alunos que utilizam ambiente de gerenciamento de aprendizado (nesse caso, o Moodle), a fim de possibilitar o aprendizado significativo, contínuo e coletivo.

Objetivos

A seguir, serão apresentados os objetivos: geral e específicos.

Objetivo Geral

O objetivo principal deste trabalho é investigar o uso de técnicas de Inteligência Artificial, no domínio da educação em computação, aplicado ao processo de ensino-aprendizagem significativo, com base no conceito de Inteligências Coletivas e sistemas de recomendação.

Objetivos Específicos

Foram definidos como objetivos específicos:

- Propor a arquitetura de um sistema especialista aplicado à educação a distância;
- Dinamizar o uso da informação com base nas Inteligências Coletivas, a fim de definir uma estratégia de ensino-aprendizagem significativa e mais eficiente;
- Proporcionar uma maior interatividade na educação a distância, oferecendo um mecanismo capaz de detectar e recomendar o conhecimento específico de cada aluno, promovendo um ambiente de compartilhamento de conhecimento;

- Oferecer um modelo de ensino virtual que utilizará uma tecnologia de aprendizado com o diferencial do auxílio da Inteligência Artificial;
- Possibilitar um ambiente capaz de promover a aquisição de conhecimentos significativos;
- Desenvolver um modelo flexível quanto aos avanços das necessidades do processo de ensino-aprendizagem.

Justificativa

A EAD se torna cada vez mais presente e, conforme os aspectos apresentados na introdução deste plano de trabalho, surge a necessidade de investigar formas de tornar o processo de ensino-aprendizagem mais eficiente com o uso das tecnologias da informação.

Conforme descrito anteriormente, existe um grande investimento das organizações privadas e governamentais no sentido de ampliar o uso da EAD no processo de ensino-aprendizado, por se caracterizar como uma solução economicamente mais viável que o modelo tradicional. O MEC (Ministério da Educação), por exemplo, tem o plano de triplicar o número de matrículas em cursos públicos de EAD até 2014, passando dos atuais 210 mil alunos para 600 mil (UOL, 2012).

Segundo João Carlos Teatini, Diretor de Educação a Distância da Capes (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), para este ano, a dotação para investimento em EAD é de R\$ 370 milhões. Ele afirma que comparar o investimento em EAD com o investimento em educação presencial é como comparar o investimento para a construção de uma rodovia com o de uma ferrovia: a ferrovia tem um investimento inicial maior, mas, para conservação, o custo é reduzido (UOL, 2012).

Filatro (2008) afirma que estudos estatísticos comprovam que, por volta de 2025, a demanda mundial por educação superior alcançará 150 milhões de pessoas contra os atuais 90 milhões de pessoas e isso não se deve, apenas, ao crescimento demográfico natural, mas também aos seguintes fatores: (i) Mudança do perfil das carreiras profissionais que enfatizam, cada vez mais, a prestação de serviços em detrimento da produção de bens e exigem competências intelectuais mais complexas, as quais necessitam de constante atualização; (ii) Mudança na política mundial, afinal, a melhoria dos índices educacionais vem, com frequência, acompanhando a democratização dos regimes políticos e vice-versa; (iii) Mudança na economia do conhecimento, uma vez que, à medida que a economia se torna mais globalizada, a atuação dos profissionais deixa de ser restrita a mercados de trabalho locais.

A ABE-EAD (Associação Brasileira dos Estudantes de Ensino a Distância) divulgou o ranking qualitativo das instituições de ensino superior particulares que oferecem cursos a distância. Essas são responsáveis pela oferta de 80,5% das matrículas em cursos superiores a distância e, em alguns casos, como marketing ou tecnologia da informação, detêm 100% dos alunos. O ranking foi

elaborado a partir de entrevistas com 16.200 estudantes de graduação a distância de 67 instituições de ensino superior (UOL EDUCAÇÃO, 2012).

Esses números concretizam a importância em se realizar pesquisa e desenvolvimento na área de EAD com a finalidade de contribuir com o crescimento e expansão da educação no Brasil.

Conteúdo do relatório

Este relatório está organizado em Fundamentação teórica, contendo a base de conhecimento proposto, Trabalhos Correlatos, uma coletânea de trabalhos similares ao desta pesquisa, Metodologia, processo utilizado para o desenvolvimento da pesquisa, desenvolvimento, o processo para o desenvolvimento do *plugin*, conclusão e as referências.

Capítulo 2

Fundamentação Teórica

Educação a Distância

A Educação a Distância é um recurso de incalculável importância como modo apropriado para atender a grandes contingentes de alunos de forma mais efetiva que outras modalidades e sem riscos de reduzir a qualidade dos serviços oferecidos em decorrência da ampliação da clientela atendida (Nunes, Ivônio Barros, 1993).

Para os que associam a educação a distância apenas a internet e acham que é um conceito recente, Walter Perry e Greville Rumble (1987) afirmam que a característica básica da educação a distância é o estabelecimento de uma comunicação de dupla via, na medida em que professor e aluno não se encontram juntos na mesma sala requisitando, assim, meios que possibilitem a comunicação entre ambos como correspondência postal, correspondência eletrônica, telefone ou telex, rádio, “modem”, videodisco controlado por computador, televisão apoiada em meios abertos de dupla comunicação, etc. Afirmam, também que há muitas denominações utilizadas correntemente para descrever a educação a distância, como: estudo aberto, educação não tradicional, estudo externo, extensão, estudo por contrato, estudo experimental.

A educação a distância apresenta várias vantagens. Em artigo publicado pelo Portal Educação - Cursos Online em 2009, Lara cita, as várias vantagens desta modalidade de ensino:

- Combinação entre estudo e trabalho.
- Permanência do aluno em seu ambiente familiar.
- Menor custo por estudante;
- Diversificação da população escolar;
- Pedagogia inovadora;
- Autonomia do aluno;
- Materiais didáticos já incluídos no preço;
- Interatividade entre alunos, professores e técnicos de apoio;
- Apoio com conteúdos digitais adicionais;
- Conteúdos desenvolvidos com orientação de aplicabilidade;
- Enfim, a EaD possibilita uma flexibilidade: Onde estudar? Quando estudar? Em que ritmo?

Mineração de Dados

No livro *Data Mining in Soft Computing Framework*, Mitra define a mineração de dados como uma forma de descoberta essencial para a resolução de problemas em um domínio específico do conhecimento. Conjuntos de dados individuais podem ser reunidos e estudados em conjunto para fins diferentes daqueles para os quais foram originalmente criados propósitos.

Novos conhecimentos podem ser obtidos no processo de eliminação, enquanto um dos maiores custos, *viz.*, A coleta de dados. Dados médicos, por exemplo, muitas vezes existe em grande quantidade em um formato não estruturado. A aplicação de mineração de dados pode facilitar a análise sistemática em tais casos.

A mineração de dados pode ser implantada de forma satisfatória em diversas áreas, a seguir Camilo e Silva (2009, apud, Witten et al. 2005, Olson et al. 2008 e Bramer 2007) apresentam algumas aplicações:

Retenção de clientes: identificação de perfis para determinados produtos, venda cruzada;

- Bancos: identificar padrões para auxiliar no gerenciamento de relacionamento com o cliente;
- Cartão de Crédito: identificar segmentos de mercado, identificar padrões de rotatividade;
- Cobrança: detecção de fraudes;
- Telemarketing: acesso facilitado aos dados do cliente;
- Eleitoral: identificação de um perfil para possíveis votantes;
- Medicina: indicação de diagnósticos mais precisos;
- Segurança: na detecção de atividades terroristas e criminais;
- Auxílio em pesquisas biométricas;
- RH: identificação de competências em currículos;
- Tomada de Decisão: filtrar as informações relevantes, fornecer indicadores de probabilidade.

Camilo e Silva (2009, apud, Ponniah 2001), também exemplifica o uso da Mineração de Dados, em:

Um supermercado na melhoria da disposição de seus produtos nas prateleiras, baseado no padrão de consumo de seus clientes;

Em uma companhia de marketing direcione o envio de mensagens promocionais, obtendo melhores retornos;

Uma empresa aérea possa diferenciar seus serviços oferecendo um atendimento personalizado;

Empresas planejem melhor a logística de distribuição dos seus produtos, prevenindo picos nas vendas;

Empresas possam economizar identificando fraudes;

E agências de viagens que possam aumentar o volume de vendas direcionando seus pacotes a clientes com aquele perfil.

Mineração de Dados Educacionais

Mineração de Dados Educacionais (do inglês, Educational Data Mining, ou EDM) é definida por Baker, Carvalho e Isotani (2011, p. 4) como a área de pesquisa que tem como principal foco o desenvolvimento de métodos para explorar conjuntos de dados coletados em ambientes educacionais. Assim, é possível compreender de forma mais eficaz e adequada os alunos, como eles aprendem, o papel do contexto na qual a aprendizagem ocorre, além de outros fatores que influenciam a aprendizagem.

Algoritmos de Similaridade

A busca por similaridade vem sendo muito empregada na área de bancos de dados, devido à necessidade de uma consulta mais inteligente. Santos Filho et al. (2001), assegura que esta técnica tem sido aplicada na área da Medicina em pesquisas por similaridade entre imagens de diagnóstico, como em mamografia, tomografia, ressonância magnética, dentre outras.

Algoritmos de pesquisa baseados nestas métricas de distância são aplicados nas bases de dados onde a consulta deve ser precisa o suficiente para que possa retornar resultados satisfatórios, e não tão precisa que venha deixar de retornar resultados relevantes.

Os métodos para associar um dado agrupamento a cada padrão seguem algum critério de similaridade e são dependentes do algoritmo empregado, dos dados utilizados e da medida de similaridade adotada.

Gestão do Conhecimento

Destaca Macedo e Matos (2010, apud, Fayyad, 1996) que a informação e o conhecimento são os elementos essenciais de qualquer organização em um mercado altamente competitivo, por meio da mineração de dados há a possibilidade de descobrir novos e úteis conhecimentos para o processo de tomada de decisão.

Consequentemente para as Instituições Acadêmicas não seria diferente, para gerar conhecimento os sistemas de informação devem ser capazes de:

- Estabelecer relações complexas entre as informações;

- Atribuir pesos diferenciados à qualidade das informações e à pertinência da aplicação das mesmas.
- Macedo e Matos (2010) concluem que a mineração de dados é uma das etapas pertencentes ao processo de KDD - Knowledge Discovery in Databases, que é composto pelas etapas de onde visa à extração de conhecimento de uma grande base de dados exigindo uma compreensão melhor das diferenças entre dado, informação e conhecimento. Por meio do conhecimento extraído este pode ser útil para as organizações, pois se trata de conhecimento gerencial fornecendo subsídios necessários para a tomada de decisão segura e confiável.

Moodle

O Modular Object-Oriented Dynamic Learning Environment é uma plataforma de aprendizagem a distância baseada em software livre, que segundo o criador do Moodle, Martin Dougiamas, trabalha com uma perspectiva dinâmica da aprendizagem em que a pedagogia socioconstrutivista e as ações colaborativas ocupam lugar de destaque.

Sabendo disso, seu objetivo permitir que processos de ensino-aprendizagem ocorram por meio não apenas da interatividade, mas, principalmente, pela interação, ou seja, privilegiando a construção/reconstrução do conhecimento, a autoria, a produção de conhecimento em colaboração com os pares e a aprendizagem significativa do aluno (SILVA, 2011).

Sistema de Recomendação

Os Sistemas de Recomendação (SR) são utilizados para identificar usuários, armazenar suas preferências e recomendar itens que podem ser produtos, serviços e/ou conteúdos, de acordo com suas necessidades e interesses.

O objetivo dos SRs não é apenas receberem o retorno de itens a partir de consultas, mas que o possível interesse do usuário por um determinado item ou serviço possa ser previsto.

Schafer (2000) analisa a estrutura dos sistemas de recomendação em quatro processos: identificação do usuário, coleta de informações, estratégias de recomendação e visualização das recomendações.

Identificação do Usuário

Muitos sistemas realizam a identificação do usuário, pois assim é possível identificar com mais precisão quem está se conectando ao sistema e assim, customizar o sistema de acordo com as características desse usuário.

Coleta de Dados

Os dados do usuário podem ser coletados de forma explícita, implícita ou inferência. Na forma explícita, as informações são coletadas do próprio usuário, por preenchimento de um formulário no site onde o mesmo informa suas preferências. Na forma implícita, as informações são coletadas pelo monitoramento da navegação do usuário com um site, páginas consultadas, histórico de compras, conteúdos selecionados, etc.

Os dados coletados são armazenados em um banco de dados e utilizados para a análise das recomendações. É possível aplicar técnicas de mineração de dados nas informações coletadas, visando descobrir relações entre as páginas visitadas, produtos comprados, conteúdos e outros itens oferecidos pelo site (Torres, 2003). A coleta de dados por inferência consiste em descobrir o perfil do usuário com base no comportamento de outros usuários com padrão de comportamento similar ao seu.

Estratégias de Recomendação

Há várias estratégias de recomendação, as mais utilizadas são: filtragem de informação, customização, sumarização estatística, individualização, *bussiness-rules*, correlação e mineração de dados. A técnica de filtragem de informação pode ser baseada em conteúdo, colaborativa ou híbrida. A Filtragem baseada em conteúdo recomenda itens semelhantes com os que o usuário demonstrou interesse no passado.

De acordo com Torres (2003) os interesses por determinados tipos de conteúdos podem definir o perfil do usuário. Na técnica de Filtragem Colaborativa é permitida a recomendação de produtos e serviços para um determinado cliente com base no que outros clientes com interesses ou preferências semelhantes tenham consumido. Diferente da filtragem de conteúdo que recomenda itens com base no conteúdo dos mesmos, a filtragem colaborativa recomenda itens que foram avaliados. A abordagem híbrida consiste do uso mútuo dessas técnicas.

Já a técnica de Customização consiste em Adaptar algo para uma determinada pessoa ou grupo de pessoas, conforme suas características. A customização é feita manualmente pelo usuário, que escolhe a maneira como a informação será apresentada. Na customização o usuário possui total controle do site Garofalakis (1999).

A Sumarização Estatística é baseada na opinião dos usuários sobre itens de um site, resumos de avaliações da comunidade e médias de popularidade. A porcentagem de usuários que compraram um determinado item ou o número de usuários que recomendam um item específico são exemplos de sumarização estatística.

A individualização é baseada na identificação de comportamentos frequentes dos usuários. Nessa recomendação o site calcula o número de visitas

a diferentes páginas e seções para coletar os interesses do usuário de forma implícita Reategui (2005) Os dados coletados são referentes ao monitoramento da navegação do usuário pelo site.

Nas recomendações business-rules (regras de negócio), o site pode utilizar informações demográficas para direcionar a propaganda de um produto para um segmento do mercado Schafer (2000), por exemplo: quando um usuário é de uma determinada região, um determinado produto é recomendado para ele, independente de seus interesses e suas necessidades.

A técnica de correlação é baseada na associação de itens que tendem a ocorrer juntamente com uma compra, podendo ser utilizados na identificação de padrões Garofalakis (1999). A correlação também pode ser definida baseada em regras de associação, que demonstra como produtos e serviços se relacionam uns com os outros.

O conhecimento extraído da mineração de dados pode ser utilizado pelos sistemas de recomendação para obter recomendações que sejam mais adequadas às preferências do usuário.

Visualização das Recomendações

As recomendações aos usuários devem ser apresentadas de maneira que as mesmas possam ser facilmente visualizadas e compreendidas. Além da apresentação das recomendações também é importante identificar os níveis de recomendação.

Segundo Schafer (2000), existem três níveis de recomendação: não recomendação, recomendação efêmera e recomendação persistente. Na não-recomendação, as recomendações são iguais para todos os usuários. Como, por exemplo, uma lista com os produtos mais vendidos para todos os seus clientes.

Na recomendação efêmera, as recomendações são baseadas inteiramente na navegação de um único usuário e não utiliza informações das navegações anteriores do mesmo. Na recomendação persistente, as recomendações são baseadas no reconhecimento do usuário, e sugere produtos que são do seu interesse, com base nas suas navegações anteriores.

Capítulo 3

Trabalhos

Correlatos

Objetos de Aprendizagem para Educação a Distância: Recursos educacionais abertos para ambientes virtuais de aprendizagem

Nesse projeto o objetivo é auxiliar, a partir de informações técnicas e pedagógicas, autores, professores e profissionais que, atuando na Educação a Distância ou presencial, necessitem criar recursos educacionais digitais, incluindo-se os objetos de aprendizagem.

Visando a facilitar a compreensão e democratização das informações que serão apresentadas, o livro identifica, apresenta e analisa softwares gratuitos que podem ser instalados no computador ou trabalhados a partir de endereços eletrônicos na Internet.

Moodle para autores e tutores - Educação a Distância na Web 2.0

Objetiva aprimorar os conhecimentos técnicos e pedagógicos de profissionais que utilizam o MOODLE como ambiente virtual de aprendizagem em seus cursos, seja a distância, seja presencial. A intenção é levar ao conhecimento dos leitores as formas de uso e potencialidades pedagógicas dos principais recursos, atividades e blocos a partir de possibilidades oferecidas pela WEB 2.0.

O livro tem como base práticas e experiências em diferentes ambientes educacionais que, efetivamente, utilizam as tecnologias de informação e comunicação digitais tanto para construção quanto para o gerenciamento e tutoria de cursos. Nesse contexto, a intenção é disponibilizar uma obra que ultrapassasse o conceito de manual, uma vez que permite ao leitor conhecer e decidir, à luz da WEB interativa, formas interessantes e criativas para comporem suas salas e ambientes virtuais.

Mineração de Dados e a Gestão Inteligente da Aprendizagem: Desafios e Direcionamentos

Esse artigo objetiva apresentar a área de mineração de dados educacionais como uma alternativa ao auxílio a esses problemas. Faremos isso tratando sobre alguns problemas clássicos e, outros, mais atuais, da educação listando técnicas que propõem solução a ele.

Aula NET

Na área educacional, a SEST SENAT¹ foca em programas de aprendizagem, que incluem preparação, treinamento, aperfeiçoamento e formação profissional, inclusive em cursos online onde o estudante acompanha as aulas a distância por meio da internet. Esse é o ambiente virtual dos Cursos on-line do SEST/SENAT.

TelEduc

O TelEduc é um ambiente para a criação, participação e administração de cursos na Web. Ele foi concebido tendo como alvo o processo de formação de professores para informática educativa, baseado na metodologia de formação contextualizada desenvolvida por pesquisadores do Unicamp.

O TelEduc foi desenvolvido de forma participativa, ou seja, todas as suas ferramentas foram idealizadas, projetadas e depuradas segundo necessidades relatadas por seus usuários. Com isso, ele apresenta características que o diferenciam dos demais ambientes para educação a distância disponível no mercado, como a facilidade de uso por pessoas não especialistas em computação, a flexibilidade quanto a como usá-lo, e um conjunto enxuto de funcionalidades.

Inteligência Coletiva Inteligente Aplicada a Sistemas Educativos

O trabalho acadêmico tem o objetivo de desenvolver a análise e concepção de um plugin para o Moodle com o intuito de fomentar a curiosidade dos alunos e dos professores a procurarem materiais de outras disciplinas, ou até mesmo alunos e professores de semestres anteriores a fim de comentarem assuntos de interesse de ambos ou até mesmo de tirarem dúvidas sobre a disciplina, como também facilitar a tomada de decisão dos coordenadores dos cursos a fim de melhorar o andamento dos alunos no curso.

Aplicação de Mineração de Dados para o Levantamento de Perfis dos Alunos do Curso de Sistemas de Informação da Facisa

O trabalho de conclusão de curso refere-se a um estudo de caso que tem como objetivo, o levantamento de perfis dos alunos do curso de Sistemas de Informação da FACISA, empregando a Mineração de Dados.

¹ SEST/SENAT: <http://www.sestsenat.org.br>

CaféEAD

A CaféEAD é uma empresa que surgiu com o intuito de oferecer as melhores soluções em EaD, transformando cada um dos projetos entregues em resultados positivos. E por intermédio da joint-venture com a Empresa WebInfo expandiram sua atuação aos mercados do Rio de Janeiro e São Paulo.

Capítulo 4

Metodologia

Este estudo pode ser classificado como pesquisa aplicada, no momento em que é proposto o desenvolvimento de um agente inteligente que possibilite recomendações entre os alunos que utilizam ambiente de gerenciamento de aprendizado (no âmbito deste trabalho, o Moodle), a fim de possibilitar o aprendizado significativo, contínuo e coletivo.

Caracterização dos objetivos da pesquisa

Esta pesquisa se classifica como exploratória, considerando que existirá a investigação do uso de técnicas de Inteligência Artificial e como essas técnicas podem contribuir com o processo de ensino-aprendizagem na computação.

Definição da natureza da pesquisa

A natureza deste trabalho será qualitativa e quantitativa. Qualitativa no momento em que sua ênfase se dará sobre aspectos de qualidade dos produtos instrucionais criados para a capacitação dos envolvidos no experimento. Quantitativo no momento em que estabelece estatísticas com base no levantamento junto às organizações a serem usadas para validação do ambiente proposto.

Caracterização do local de realização da pesquisa

Este trabalho se caracteriza como uma pesquisa de campo, uma vez que ela se dará no ambiente natural, ou seja, em organizações que adotam Sistemas de Gestão de Aprendizado (em especial, o Moodle).

Definição das técnicas de pesquisa

As técnicas a serem aplicadas buscarão garantir o alcance dos objetivos específicos e, conseqüentemente, do objetivo geral, que se adequarão a cada um dos objetivos estabelecidos:

Caracterização e análise as técnicas para criação de material instrucional

- Levantamento bibliográfico sobre a produção de materiais instrucionais dando ênfase ao ensino-aprendizado da computação;
- Levantamento, junto a organizações que já implementaram ou estão em processo de implementação de Sistemas de Gestão de Aprendizado sobre os processos e critérios por elas adotados no contexto de implementação desses. Esse levantamento permitirá uma validação dos processos e critérios obtidos no levantamento bibliográfico.
- Caracterização das áreas de domínio da Inteligência Artificial
- Levantamento bibliográfico sobre áreas de domínio da Inteligência Artificial no contexto educacional;
- Estabelecimento das associações entre áreas de domínio da Inteligência Artificial, Produção de Material Instrucional, Mineração de Dados Educacionais, Educação a Distância e Sistemas de Recomendação;
- Desenvolvimento do Agente Inteligente baseado nas análises técnicas das diversas áreas estudadas;
- Validação da ferramenta desenvolvida, análise dos dados, realização de melhorias e, por fim e produção de artefatos científicos, a fim de contribuir com a área estudada.

Capítulo 5

Desenvolvimento

A plataforma Moodle tem uma excelente biblioteca para desenvolvimento e extensão de suas funcionalidades. Trata-se de um conjunto de funções que compõe um módulo central do sistema. Essas funções são utilizadas para implementar os módulos que compõem toda a plataforma.

O trabalho de desenvolvimento e customização do Moodle se torna muito mais rápido e fácil com o uso dessa biblioteca e toda a biblioteca fica organizada na pasta lib no endereço raiz da instalação do Moodle.

Para torná-la disponível basta incluir o arquivo *config.php* no início do arquivo de programação php com o seguinte comando:

```
require_once("config.php");
```

Criação do Módulo Agente de Recomendação

Para a criação de um novo módulo para a plataforma Moodle é necessária a utilização de uma estrutura pré-definida de criação. Todos os itens dessa estrutura são essenciais para que o novo módulo seja reconhecido pela plataforma.

Essa estrutura consiste em diretórios, arquivos de instalação e configuração, arquivos de internacionalização, imagens e os arquivos com as funcionalidades específicas do módulo, conforme visualizado na figura 1 e especificado na tabela 1, abaixo.

Name ^	Date modified	Type	Size
db	25/09/2007 11:06	File folder	
lang	20/01/2013 08:04	File folder	
icon.gif	30/09/2003 10:45	GIF Image	1 KB
index.php	01/09/2009 08:12	PHP File	3 KB
lib.php	23/04/2009 08:13	PHP File	6 KB
mod_form.php	20/03/2009 08:11	PHP File	4 KB
README.txt	19/03/2009 08:11	Text Document	4 KB
version.php	20/03/2009 08:11	PHP File	1 KB
view.php	18/04/2009 08:12	PHP File	3 KB

Figura 1: Disponibilização dos Arquivos do Módulo de Agente Inteligente Educacional

Tabela 1: Descrição do módulo

Item	Descrição
db/install.xml	Descreve as tabelas necessárias para o módulo.
db/upgrade.php	Utilizado para alterar a estrutura das tabelas do seu módulo (quando necessário)
icon.gif	Ícone que representa a logomarca do Agente
index.php	Lista todas as instâncias da funcionalidade do módulo
lang/en_utf8/help/newmodule/index.html	Utilizado para definição do idioma
lang/en_utf8/help/newmodule/mods.html	Utilizado para definição do idioma
lang/en_utf8/newmodule.php	Utilizado para definição do idioma
lib.php	Funções base necessárias para o módulo
mod_form.php	Descreve a forma da instância de módulo
README.txt	Informações sobre o funcionamento do módulo
version.php	Arquivo de controle de versão do módulo
view.php	Utilizado para a primeira execução do módulo

Tecnologias Utilizadas

No desenvolvimento do *plugin* foram utilizadas as seguintes tecnologias:

MySQL Workbench

Trata-se de uma ferramenta visual multiplataforma para arquitetos de banco de dados, programadores e DBAs. Fornece modelagem de dados, desenvolvimento de SQL e ferramentas para administração de configuração.

O Workbench foi utilizado na modelagem do banco de dados e criação do modelo entidade-relacional. A figura 2 apresenta a tela inicial do MySQL.

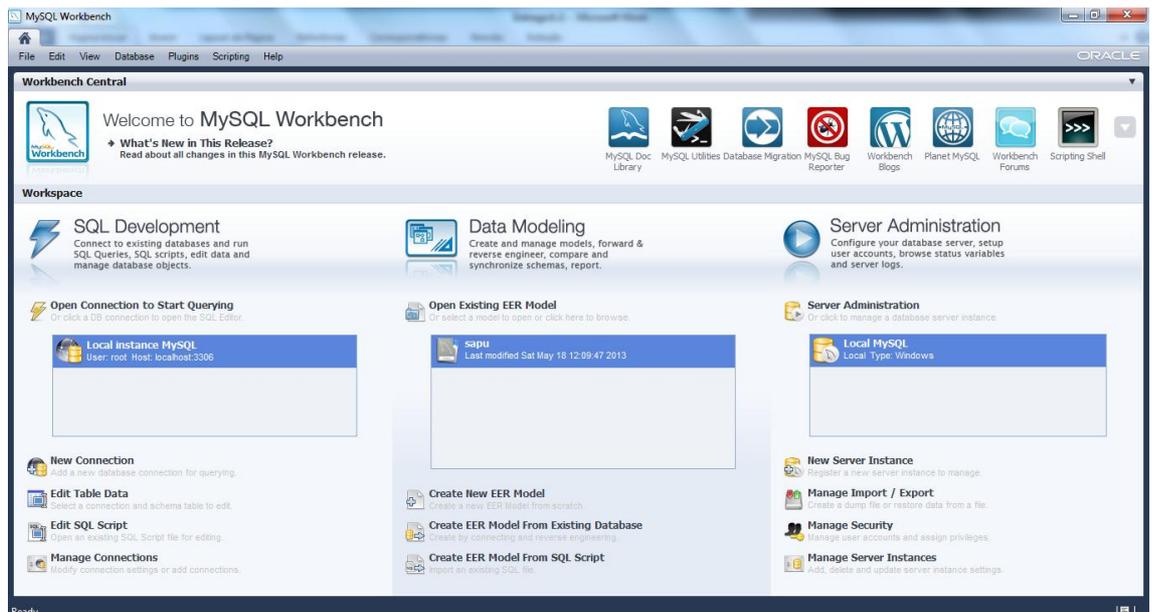


Figura 2 - Tela inicial do MySQL Workbench
Fonte: Autor

PHP

A linguagem PHP se caracteriza como uma linguagem de scripting (interpretada). Ela é amplamente utilizada para o desenvolvimento web podendo ser incorporada ao HTML. Foi adicionado ao projeto principalmente para fornecer o acesso aos dados no banco e porque o Moodle é desenvolvido em PHP.

Função de Similaridade

Foi desenvolvido um módulo para cálculo simples de similaridade. Este módulo recebe os perfis dos usuários e realiza o cálculo de similaridade utilizando a distância euclidiana. Na figura 3 apresenta o código utilizado para este cálculo.

```
<?
// retorna a porcentagem da similaridade entre dois textos
function str_compare($str1, $str2) {
    $count = 0;

    $str1 = ereg_replace("[^a-z]", ' ', strtolower($str1));
    while(strpos($str1, ' ')) {
        $str1 = str_replace(' ', ' ', $str1);
    }
    $str1 = explode(' ', $str1);

    $str2 = ereg_replace("[^a-z]", ' ', strtolower($str2));
    while(strpos($str2, ' ')) {
        $str2 = str_replace(' ', ' ', $str2);
    }
    $str2 = explode(' ', $str2);

    if(count($str1)<count($str2)) {
        $tmp = $str1;
        $str1 = $str2;
        $str2 = $tmp;
        unset($tmp);
    }

    for($i=0; $i<count($str1); $i++) {
        if(in_array($str1[$i], $str2)) {
            $count++;
        }
    }

    return $count/count($str2)*100;
}
```

Figura 3: Módulo de Cálculo de Similaridade

Conclusão

A pesquisa “Análise e Desenvolvimento de Métodos para Mineração de Dados Educacionais” teve como objetivo a análise de algoritmos, recomendações e educação.

Além da análise sobre o estado-da-arte, foi proposto o desenvolvimento de um plugin para a plataforma Moodle que utilizasse dos dados dos alunos cadastrados na plataforma e, com o uso de algoritmos de inteligência artificial e algoritmos de recomendação, fosse possível o cálculo de similaridade entre os perfis dos usuários, culminando com a recomendação de conteúdo.

Na versão atual do projeto, o plugin foi desenvolvido. A próxima atividade está relacionada à execução de um caso de estudo, a fim de validar o desenvolvimento do *plugin* e das técnicas de recomendações utilizadas no processo de desenvolvimento.

Um artigo científico sobre os resultados desta pesquisa também será escrito e submetido para o SBSI 2015 (Simpósio Brasileiro em Sistemas de Informação).

O produto final, após a execução do estudo de caso, será disponibilizado no ambiente Github (Serviço de Web Hosting Compartilhado) para que outros pesquisadores possam utilizar do plugin em suas pesquisas.

Aspectos diretos, tais como o processo de iniciação científica de quatro (4) alunos de graduação, bem como a continuação deste projeto pelo desenvolvimento de um trabalho de conclusão de curso – TCC.

Referências

Ambiente Virtual de Aprendizagem ou Learning Management System. Disponível em: <<http://cafeead.com.br/site/solucoes/lms>>. Acesso em 16maio 2013.

BAKER, R. S. J. D., ISOTANI, S., DE CARVALHO, A., **Mineração de Dados Educacionais: Oportunidades para o Brasil**. Revista Brasileira de Informática na Educação, 2011.

BAKER, Ryan Shaun Joazeiro de; ISOTANI, Seiji; CARVALHO, Adriana Maria Joazeiro Baker de. **Mineração de Dados Educacionais: Oportunidades para o Brasil**.

BORGES, Eduardo N; CONY, Carlos A; FREITAS, André L. C. de. **Utilização de Algoritmos de Busca por Similaridade em SGBDs Comerciais**. Rio Grande-RS, 2005.

CAMILO, Cássio Oliveira; SILVA, João Carlos da. **Mineração de Dados: Conceitos, Tarefas, Métodos e Ferramentas**. Universidade Federal de Goiás, 2009.

CARVALHO, L. V., Trabalho de conclusão de curso que será publicado no final deste ano sobre Inteligência Coletiva Inteligente no curso de Sistemas de Informação da FACISA, 2012.

CARVALHO, Liliane Vilar de. **Inteligência Coletiva Inteligente Aplicada a Sistemas Educativos**. Campina Grande-PB, 2012.

Censo EAD.BR: **relatório analítico da aprendizagem a distância no Brasil 2011**. São Paulo: Pearson Education do Brasil, 2012.

Cursos online SEST SENAT. Disponível em: <<http://aulanet.sestsenat.org.br/>>. Acesso em 03maio 2013.

DA SILVA, Robson Santos. **Moodle para Autores e Tutores: Educação a Distância na Web 2.0**. Segunda Edição. São Paulo. Novatec, 2011.

DANIELA, R. Silva; MARINA, T. P. VIEIRA; Wesley, P. Seno. **Acompanhamento do Aprendizado em Educação a Distância com Uso de Data Mining**. Universidade Federal de São Carlos, São Paulo.

DAVID, Eduardo Gonçalves. **IA na logística da EAD**. In: ABED. Disponível em <<http://www.abed.org.br/congresso2005/por/pdf/143tce4.pdf>>. Acesso em 02 de novembro de 2012.

Ensino a distância: uma realidade crescente no Brasil, 2009. Disponível em: <http://www.abed.org.br/site/pt/midiateca/clipping_abed/909/2009/07/ensino_a_distancia_uma_realidade_crescente_no_brasil_>. Acesso em: 04 maio 2013.

FILATRO, Andrea. **Design Instrucional na Prática**. In: Prentice Hall Brasil, 2008.

FIORAVANTI, M. L.; NAKAGAWA, E. Y. ; BARBOSA, E. F. **EDUCAR: Uma Arquitetura de Referência para Ambientes Educacionais**. In: XXI Simpósio Brasileiro de Informática na Educação (SBIE 2010), João Pessoa (PB). XXI Simpósio Brasileiro de Informática na Educação (SBIE 2010), 2010.

FORMIGA, Marcos. LITTO, Fredric M. **Educação a Distância - O Estado da Arte**. Prentice Hall Brasil, 2009.

GIRAFFA, Lucia M. M.; SOUZA, Rafael V. de; MARCZAK, Sabrina; PRIKLANDNICKI, Rafael. **PDS-E: Em direção a um processo para desenvolvimento de Software Educacional**. In: Anais do XXV Congresso da SBC, São Leopoldo – RS. Disponível em: <www.brie.org/pub/index.php/wie/article/view/867/853>. Acesso em 01 novembro 2012.

HOLDEREGGER, Ricardo et al. **O uso de agentes de conhecimento na construção do novo EAD para o ensino superior tecnológico**. Workshop de Pós-Graduação e Pesquisa - CPS. São Paulo – SP, 2008.

ISOTANI, S. ; BITTENCOURT, I. ; JAQUES, P. **Mineração de Dados e a Gestão Inteligente da Aprendizagem: Desafios e Direcionamento**. In: CSBC 2012 - Desafie, 2012, Curitiba. Anais do CSBC. Porto Alegre: SBC, 2012.

JAIED - International Journal of Artificial Intelligence in Education. In: IOS Press. Adquirido em: <<http://www.iospress.nl/journal/international-journal-of-artificial-intelligence-in-education/>>, 2012.

KAMPFF, A. J. C. **Mineração de dados educacionais para geração de alertas em ambientes virtuais de aprendizagem como apoio à prática docente**. In: Tese de doutorado em Informática na Educação. Universidade Federal do Rio Grande do Sul, 2009.

LARA, Enderson, EAD – **Vantagens da Educação a Distância**. Disponível em: <<http://www.portaleducacao.com.br/pedagogia/artigos/7671/ead-vantagens-da-educacao-a-distancia>>. Acesso em: 05 maio 2013.

LÉVY, P. **A inteligência coletiva por uma antropologia do ciberespaço**. In: Edições Loyola, São Paulo, 1998.

LÉVY, P.. **Pierre Lévy destaca relação entre a educação e a era digital**.

Disponível em:

<<http://redeglobo.globo.com/globouniversidade/noticia/2012/08/pierre-levy-destaca-relacao-entre-educacao-e-era-digital.html>>. Acessado em: 02 novembro de 2012

LOOI, Chee-Kit. MCCALLA, Gord. BREDEWEG, Bert. BREUKER, Joost. **Artificial Intelligence in Education: Supporting Learning through Intelligent and Socially Informed Technology**. In: IOS Press, 2005.

LUZZI, Daniel; LUSWARGHI, Andrea. **Os Desafios da Educação a Distância no Contexto Latino-Americano**, 2005. Disponível em: <http://www.abed.org.br/site/pt/midiateca/textos_ead/687/2005/11/os_desafios_da_educacao_a_distancia_no_contexto_latino-americano_>. Acesso em: 04 maio 2013.

MACEDO, Dayana Carla de; MATOS, Simone Nasser. **Extração de Conhecimento Através da Mineração de Dados**. 2010.

MAIA, Carmem. NETO, Joao Augusto Mattar. **ABC da EAD: a educação a distância**. Prentice Hall Brasil, 2008.

MEC - Secretaria de Educação a Distância.. Disponível em: <http://portal.mec.gov.br/index.php?option=com_content&view=article&id=12502&Itemid=823>. Acessado em 04 novembro de 2012.

MOODLE. **Moodle**. Disponível em: <<http://moodle.org/>>. Acessado em: 04 novembro de 2012.

MORAN, José Manuel. **Tendências da Educação online no Brasil**. Rio de Janeiro. Qualitymark, 2005.

MOREIRA, Mireille Pinheiro. FAVERO, Eloi Luiz. **Um Ambiente para Ensino de Programação com Feedback Automático de Exercícios**. In: [WEI] - Anais do Workshop sobre Educação em Informática. Bento Gonçalves – RS, 2009.

MOURSUND, D.G.. **Brief introduction to educational implications of Artificial Intelligence**. Disponível em <<http://darkwing.uoregon.edu/~moursund/>>. Acesso em 03 novembro 2012.

OGAN, Amy. *at al.* **Collaboration in Cognitive Tutor Use in Latin America: Field Study and Design Recommendations**. ACM. Austin. TX – USA, 2012.

PAIVA, Ranilson; BITTENCOURT, Ig Ibert; PACHECO, Henrique; SILVA, Alan Pedro da; JACQUES, Patrícia; ISOTANI, Seiji. **Mineração de Dados e a Gestão Inteligente da Aprendizagem**: Desafios e Direcionamentos.

PAVLIK, P., CEN, H., Wu, L. KOEDINGER, K. **Using Item-type Performance Covariance to Improve the Skill Model of an Existing Tutor**. In Proceedings of the International Conference on Educational Data Mining, 2008, 77-86.

PIMENTEL, A R., DIRENE A I. **Medidas Cognitivas no Ensino de Programação de Computadores com Sistemas Tutores Inteligentes**. In: Simpósio Brasileiro de Informática Educativa, SBIE'98, Fortaleza - CE, Brasil, 1998.

REZENDE, Solange Oliveira. **Mineração de Dados**, 2003.

ROESLER, Jucimara. **As Redes Sociais e a Comunicação na Educação On-Line**. In: Revista Linhas - Revista do Programa de Pós-Graduação em Educação - V12. Florianópolis, 2011.

ROMERO, C., VENTURA, S., GARCIA, E. **Data mining in course management systems: Moodle case study and tutorial**. Computers & Education, 2008, 368–384.

SANTOS FILHO, R. F.; SOUSA, E. P. M. de; TRAINA, A. J. M.; TRAINA, Júnior. **Desmistificando o Conceito de Consultas por Similaridade: A Busca de Novas Aplicações na Medicina**. WIM 2001.

SANTOS, Adriano Araújo *et al.* **Logic SORT: Uma Oportunidade de Aprendizagem Alternativa sobre Algoritmos de Ordenação**. In: CSBC 2012 - [WEI] Anais do Workshop sobre Educação em Informática. Curitiba – Paraná, 2012.

SCHEUER, O. MCLAREN, B.M. **Helping teachers handle the flood of data in online student discussions**. In Proceedings of the International Conference on Intelligent Tutoring Systems (ITS-08). 2008, 323-332.

SILVA, Robson Santos da. **Moodle para autores e tutores. 2ª ed. São Paulo: Novatec, 2011.**

SILVA, Robson Santos da. **Objetos de Aprendizagem para Educação a Distância - Recursos educacionais abertos para ambientes virtuais de aprendizagem. São Paulo: Novatec, 2011.**

SOUTO, Lizianne Priscila Marques. **Aplicação de Mineração de Dados para o Levantamento de Perfis dos Alunos do Curso de Sistemas de Informação da Facisa**. Campina Grande-PB, 2012.

Teleduc Educação a Distância. Disponível em: <<http://www.teleduc.org.br/>>. Acesso em 03maio 2013.

UOL Educação. **MEC promete triplicar matrículas em EAD e alcançar 600 mil alunos até 2014**. Disponível em: <<http://educacao.uol.com.br/noticias/2012/04/24/mec-promete-triplicar-matriculas-em-EAD-ate-2014-e-alcancar-600-mil-alunos.htm>>. Acessado em 02 novembro de 2012.

UOL Educação. **Ranking de qualificação da ABE-EAD (Associação Brasileira dos Estudantes de Ensino a Distância) das faculdades particulares que oferecem cursos a distância**. Disponível em: <<http://educacao.uol.com.br/noticias/2012/06/05/confira-o-ranking-das-faculdades-particulares-que-oferecem-cursos-a-distancia.htm>>. Acessado em novembro 2012.