

Revista Científica da Escola de Administração do Exército

Volume 1 - Nº 1 - 1º semestre de 2005
ISSN: 1808 - 5784



Revista Científica da Escola de Administração do Exército

Vol 1 – nº 1 - 1º semestre de 2005

ISSN: 1808 – 5784



Escola de Administração do Exército – EsAEx

Comandante: Tenente-Coronel Lourival Carvalho Silva

Sub Comandante: Tenente-Coronel Carlos Cordeiro da Silva Santos

Chefe da Divisão de Ensino: Tenente-Coronel José Lindolfo Weber da Silva

Conselho Editorial

Tenente-Coronel Marcos Nalin

Capitão Carla Christina Passos

Capitão José Roberto Pinho de Andrade Lima

Capitão Marçal de Lima Hokama

Capa e Editoração Eletrônica: Capitão Marçal de Lima Hokama

Redação: caplima@esaex.mil.br

Revista digital disponível no site da Escola de Administração do Exército

<http://www.esaex.ensino.eb.br>

Escola de Administração do Exército

Rua Território do Amapá, 455 – Pituba

Salvador - BA

CEP: 41830-540

Revista Científica da Escola de Administração do Exército. – Vol. 1, nº 1 (1º semestre de 2005)- . – Salvador: Escola de Administração do Exército, 2005.
v. : il.

Publicação Semestral
ISSN 1808-5784

1. Tecnologia. 2. Educação. 3. Saúde. I. Escola de Administração do Exército.

Todos os direitos reservados. Proibida a reprodução sem autorização prévia e escrita.

Todas as informações técnicas são de responsabilidade dos respectivos autores.

Índice

Editorial

Tecnologia, 5

- A Descoberta do Conhecimento como Ferramenta Estratégica para o Exército Brasileiro, 6
- Comparativo do Sistema de Cadastro de Alunos do Exército desenvolvido na EsAEx à luz do nível 2 de maturidade do CMMI-SW, 17
- Avaliação Construtivista Informatizada no Ensino a Distância: Uma Proposta para o Colégio Militar de Salvador, 28
- Redes Neurais no Auxílio Pedagógico aos Cursos de Formação Militar, 39
- Uma proposta para a implantação de um ambiente de desenvolvimento de software com segurança, 52
- Segurança da Informação: Ameaças e Vulnerabilidades de uma OM do EB, 70

Educação, 84

- Tecnologia da Educação na perspectiva de uma Gestão de Qualidade, 85
- Causas mais frequentes das dificuldades de aprendizagem: Uma proposta de identificação e treinamento para os Colégios Militares, 100
- Contribuições Teóricas para o Aperfeiçoamento da Liderança Militar no Exército Brasileiro, 113

Saúde, 124

- Atividade física relacionada à Qualidade de Vida, 125
- Proposta de um Indicador de Qualidade para os Serviços de Aprovisionamento do Exército (IQSA), 137
- Análise dos Indicadores Epidemiológicos dos Oficiais Alunos do Curso de Formação de Oficiais do Quadro Complementar do Exército: Uma Proposta de Educação em Saúde, 155

Editorial

A **Revista Científica da Escola de Administração do Exército** (EsAEx) surge como um espaço, há muito esperado, para a publicação de trabalhos científicos desta escola, onde são formados os oficiais do Quadro Complementar do Exército Brasileiro. Compõem este primeiro volume os doze artigos científicos destaques do Curso de Formação de Oficiais do ano de 2004, com temas ligados à Informática, Administração, Vigilância Sanitária, Epidemiologia, liderança e ferramentas educacionais.

Esta revista busca dar visibilidade ao produto intelectual dos integrantes da linha do ensino complementar do Exército. Portanto, é o local ideal para a publicação dos trabalhos científicos de todos os integrantes do Quadro Complementar de Oficiais (QCO), formados desde 1990, primeira turma da EsAEx. Muitos militares, sabemos, concluíram especializações, mestrados e doutorados, e têm aqui a oportunidade de dividir o conhecimento auferido com os demais interessados na investigação científica e na produção de saber.

A **Revista Científica da Escola de Administração do Exército** está aberta a todos os trabalhos científicos desenvolvidos nos diversos ramos da linha de ensino complementar, ou seja: ciências administrativas (Administração, Economia, Contabilidade e Estatística), Informática, ciências humanas (Direito, Comunicação Social e Psicologia), magistério (inclusive Pedagogia) e Ciências Agrárias e da saúde (veterinária e enfermagem).

Concitemos, assim, a todos que produzem conhecimento nos campos supracitados, a construir, junto com a equipe editorial da **Revista Científica da Escola de Administração do Exército**, um instrumento dinâmico, atual, criativo, inovador, porém, em estreita ligação com a realidade, com capacidade de contribuir para a modernização da Força Terrestre.

LOURIVAL CARVALHO SILVA – Tenente-Coronel
Comandante da Escola de Administração do Exército

Tecnologia

A Descoberta do Conhecimento como Ferramenta Estratégica para o Exército Brasileiro

LUCIANO GARCIA MOREIRA¹, ADRIANO LUIZ DE OLIVEIRA GOMES², MARCOS ROBERTO MARCONDES³

Resumo. Este artigo procura mostrar a possibilidade de uso estratégico da Descoberta do Conhecimento pelo Exército Brasileiro em suas bases de dados, bem como situações do mundo real as quais foram desenvolvidas, baseadas nos estudos científicos realizados. Também é feita uma explicação do funcionamento do processo da Descoberta do Conhecimento denominado KDD (Knowledge Discovery in Databases) e sua principal etapa chamada Mineração de Dados/Data Mining.

Palavras-chave: Descoberta do Conhecimento, Data Warehouse, Mineração de Dados, Data Mining, Banco de Dados.

Abstract. This article aims to show the strategic use of Knowledge Discovery by the Brazilian Army in its databases as well as realistic situations that were developed based upon previous scientific studies. There is also an explanation of the Knowledge Discovery process called KDD and its main stage called Data Mining.

Key-words: Knowledge Discovery, Data Warehouse, Data Mining, Databases.

1 Introdução

A informação é o bem mais precioso de qualquer Organização (SILVA, 2000, p. 2). Cada vez mais a obtenção do conhecimento e de informações estratégicas são necessárias para a tomada de decisões coerentes e precisas. O problema a ser abordado neste trabalho é que as Organizações se tornaram verdadeiras mestras em armazenar dados, no entanto se mostram ineficientes em aproveitá-los e/ou analisá-los de forma a adquirir o conhecimento.

Para contextualizar este problema e propor uma solução, será considerado o processo de Descoberta/Prospecção do Conhecimento, isto é, a procura de relações escondidas em um grande conjunto de dados. Este processo tornou-se tema de estudos científicos a partir dos anos noventa, quando começou a surgir o consenso de que a imensa massa de dados era, na verdade, um grande tesouro e precisava ser muito bem explorada. De acordo com Schneider (2002, p. 2), o termo Descoberta do Conhecimento em

¹ Tenente-Aluno do Curso de Formação de Oficiais do Quadro Complementar de 2004. Graduado em Informática. lucgar@ig.com.br.

² Tenente-Aluno do Curso de Formação de Oficiais do Quadro Complementar de 2004. Graduado em Informática. mralog@yahoo.com.br.

³ Capitão do Quadro Complementar de Oficiais. Mestre em Modelagem Computacional. marcos_marcondes@uol.com.br.

Base de Dados (*Knowledge Discovery in Databases*) foi formalizado em 1989, em referência ao processo de procurar conhecimento em dados, e se propõe a descobrir novas relações, padrões e tendências significativas, por meio da análise precisa de grande conjunto de dados gravados. Tais dados válidos, potencialmente úteis e interpretáveis vêm a se tornar a tão desejada informação estratégica, de grande valia e relevância para a tomada de decisão (SANTOS, 2002, p. 3).

Propõe-se, por meio deste trabalho, um estudo criterioso sobre o processo da Descoberta do Conhecimento, identificando, conceituando e demonstrando suas fases e métodos, buscando associar a importância deste processo e a utilização do mesmo em bases de dados do Exército Brasileiro.

O estudo está dividido da seguinte forma: A seção 2 explicará o processo da Descoberta do Conhecimento (*KDD*); A seção 3 descreverá o conceito de Mineração de dados, que é a principal etapa do processo, e sua relação com outras áreas. Em seguida, a seção 4 apresentará as Tarefas, Técnicas e Algoritmos da Mineração. Já a seção 5 demonstrará as áreas de aplicação e evidências da Descoberta do Conhecimento. A seção 6 mostrará uma proposta de utilização do processo nas bases de dados do Exército Brasileiro.

2 Descoberta de Conhecimento em Banco de Dados (*Knowledge Discovery in Databases – KDD*) - DCBD

O processo *KDD* tem o objetivo de identificar dados que sejam informações válidas, estratégicas e compreensíveis.

O processo do *KDD* é composto por etapas seqüenciais, possuindo uma natureza iterativa e interativa

(SCHNEIDER, 2002, p. 2). É considerado iterativo pois, mesmo as etapas sendo seqüenciais, poderá haver retorno a alguma destas. Enfatiza-se a interatividade, já que o usuário poderá solicitar retorno a alguma das etapas do processo, caso haja necessidade de ajustes e correções. Conforme Oliveira (2000, p. 13), as pessoas envolvidas na sua realização devem possuir um canal de comunicação que viabilize uma boa troca de informações. Assim, pode-se destacar diferentes especialidades destas pessoas, como: especialista do domínio, que é o que possui ampla experiência da situação proposta; analista, que é o que possui amplo conhecimento das etapas do *KDD* e o usuário final, o qual receberá e trabalhará com os dados obtidos para tomada de decisões. Tais especialistas identificam os problemas, descrevem os objetivos e projetam a arquitetura. Já os computadores têm a incumbência de processar os dados na procura de padrões a fim de satisfazer os profissionais envolvidos no processo.

Conforme Figura 1, o processo de Descoberta do Conhecimento – *KDD* está dividido nas seguintes fases:

- a) Definição dos Objetivos: fase pela qual define-se os objetivos e requisitos necessários à extração do conhecimento, as possíveis fontes de dados e o estudo de viabilidade e custos da aplicação;
- b) Seleção: nesta etapa, acontece o agrupamento dos dados, isto é, reunião de todos os dados necessários à prospecção/extração do conhecimento;
- c) Pré-Processamento e Limpeza: nesta situação, temos a integração de dados heterogêneos, eliminação de incompatibilidade de dados e registros duplicados. Esta etapa pode até tomar 80

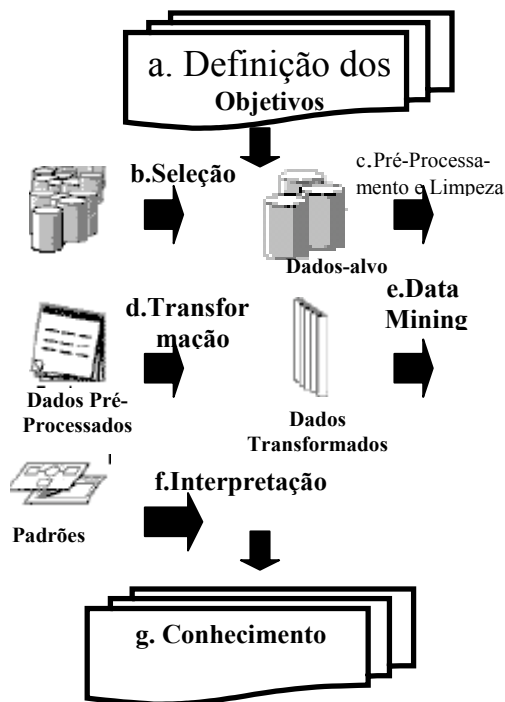


Figura 1: Esquema do Processo da Descoberta do Conhecimento

% de todo o processo, devido às bem conhecidas dificuldades de integração de dados heterogêneos (FIGUEIRA,1998, p. 8).

A fim de diminuir o tempo de execução da Descoberta do Conhecimento, se uma Organização já possui um *data warehouse*, as tarefas citadas nas letras b e c não precisam ser executadas, visto que, com a tecnologia descrita, os dados estão prontos para a próxima fase. Segundo Come (2001, p. 21), o *data warehouse* teve origem a partir dos estudos do MIT (*Massachusetts Institute of Technology*), nos anos 70, e pode ser definido como um repositório de dados centralizado, ou seja, os dados já se encontram organizados, resultante da integração e prontos para serem manipulados. Com o repositório, os dados são apenas disponibilizados e consultados, não é necessário confrontar várias aplicações em busca de uma inteligência e administração nos negócios. A Figura 2 mostra um exemplo de *data*

warehouse. Supõe-se os bancos de dados de várias áreas de uma Instituição/Empresa ligados a um *data warehouse*.

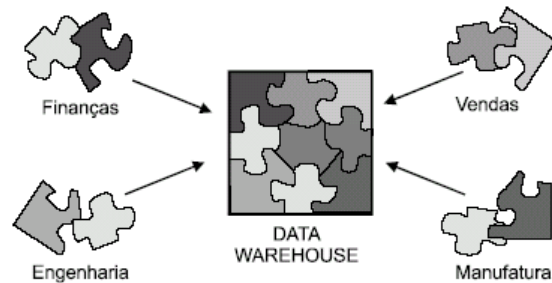


Figura 2:Exemplo de um *Data Warehouse* (COME, 2001, p. 2)

Dando continuidade ao estudo das etapas, apresenta-se:

d) Transformação: a fim de facilitar a etapa do *Data Mining*, os dados pré-processados ainda passam por uma transformação, onde são reduzidos para diminuição de variáveis envolvidas. Esta transformação busca atributos que representem todas as características destes dados (SCHNEIDER, 2002, p. 3);

e) Mineração de Dados/*Data Mining*: tal processo envolve a busca por padrões de dados, através da escolha de técnicas e algoritmos próprios;

f) Interpretação: com a mineração de dados aplicada, é feita uma avaliação com o objetivo de ponderar novas descobertas e pode ser necessário o retorno a alguma das fases anteriores;

g) Conhecimento: após a validação do conhecimento, este deve servir a um sistema de apoio à tomada de decisão, através de técnicas de representação e visualização como tabelas, gráficos e grafos.

Nos estudos realizados, nota-se que a última fase é desconsiderada por alguns autores. De acordo com a análise feita, conclui-se que é uma etapa a considerar,

pois é quando todo o trabalho desenvolvido é apresentado, visando a tomada de decisões futuras que serão estratégicas e racionais.

Para que o conhecimento extraído seja realmente utilizado, todos os usuários que fazem parte do processo decisório devem receber todas as informações sobre novas descobertas encontradas (OLIVEIRA, 2000, p. 25).

É importante enfatizar que utiliza-se com frequência o termo *Data Mining* / Mineração de Dados como sinônimo do processo do *KDD*. Assim, mostra-se que este é a Descoberta do Conhecimento como um todo e a Mineração de Dados é uma das etapas do referido processo.

3 Mineração de dados e sua Relação com Outras Áreas

A Mineração de Dados é o núcleo do *KDD* e a sua mais importante etapa. Devido a esta situação, a Mineração é considerada um processo único, principalmente na comunidade de estatísticos e analistas de sistemas. No âmbito da pesquisa de Inteligência Artificial, o termo usado já é o correto, isto é, *KDD*. Assim, existe a definição como sendo um termo genérico utilizado para todas as técnicas computacionais para extração de informações úteis, a partir de grandes conjuntos de dados armazenados (SCHNEIDER, 2002, p. 3).

As ferramentas de mineração de dados têm o papel desbravador de descobrir, previamente, características dos dados, sejam relacionamentos, dependências ou tendências, procurando anomalias e possíveis relações, identificando, assim, problemas que não tinham sido identificados antes, fazendo uso de algoritmos próprios. Os resultados obtidos podem ser usados no gerenciamento de informação, processamento de pedidos de informação,

tomada de decisão, controle de processo e muitas outras aplicações.

O propósito da Mineração de Dados é a simplificação dos dados. É impossível descobrir conhecimento sem desconsiderar o específico e valorizar o genérico. Eventos isolados podem não conter informação valiosa, podem não indicar nenhuma tendência, apenas uma necessidade específica naquele determinado momento. É fundamental saber que a Mineração de Dados não se preocupa em entender o *porquê* da presença de determinado dado em um banco de dados, mas sim a consequência da presença deste dado.

Mineração de Dados é uma área de interseção de Estatística, Aprendizado de Máquina e Banco de Dados (SCHNEIDER, 2002, p. 3). Mineração e Estatística são tecnologias distintas, mas apresentam pontos em comum. Por exemplo, a Mineração utiliza conceitos da Estatística para produzir seus resultados. Outras comparações podem ser feitas, como por exemplo, a função principal da Estatística é o modelo, da Mineração é o algoritmo. A Mineração pode ser realizada em cima de dados não estruturados como vídeo e imagens. O conjunto de dados usados na Mineração é muito volumoso para a Estatística; e o uso da Mineração é para usuários comuns, não para estatísticos, já que ela automatiza o processo da Estatística. O que pode ser dito em relação à Mineração e ao Aprendizado de Máquina é que a primeira faz parte do Aprendizado de Máquina, sendo que este está voltado para melhorar a performance de um agente e aquela para encontrar um conhecimento compreensível. Outra diferença é que a Mineração trabalha com grandes volumes de dados e o Aprendizado, com volumes menores. A relação que se percebe entre Mineração de Dados e Bancos de Dados é

clara: a Mineração deve ser aplicada em cima de um conjunto de dados de um Banco de Dados. Embora possa ser feita em cima de uma base de dados de um sistema transacional, ou seja, com dados brutos sem preparação para a Descoberta, seria muito mais fácil se os dados já estivessem organizados em um *data warehouse*, tornando o dispêndio tecnológico e computacional bem menor e obtendo um ganho estratégico nos resultados muito maior.

4 Tarefas, Técnicas e Algoritmos de Mineração de Dados

Para aplicar o processo de *KDD*, e em consequência, a Mineração de Dados em alguma aplicação, devemos levar em conta que existem diversas tarefas que podem ser utilizadas e, associadas a elas, outras tantas técnicas de Mineração de Dados (SCHNEIDER, 2002, p. 4).

Pode-se entender uma tarefa como os relacionamentos que são estabelecidos entre os dados para se obter conhecimento interpretável. As tarefas são divididas em quatro categorias, como visto na Tabela 1: agrupamento, associação, classificação e correção de desvios. Agrupamento é uma tarefa que junta os registros heterogêneos de dados em segmentos homogêneos, de acordo com alguma similaridade entre eles. É realizada, normalmente, como antecessora de alguma outra técnica. Já a tarefa de associação identifica registros de dados relacionados, ou seja, encontra elementos que indiquem a presença de outros, revelando afinidades escondidas entre eles (BARIONI, 2002, p. 22). Uma variação desta tarefa está nos padrões sequenciais, que procuram eventos que estejam relacionados através do tempo. Por exemplo, pode-se descobrir que a aquisição de um produto geralmente implica, algum tempo depois, na

aquisição de um outro produto. Por outro lado, a tarefa de classificação consiste em construir um modelo de algum tipo que possa ser aplicado a dados não classificados, visando categorizá-los em classes. Esta tarefa pode ser considerada uma tarefa indeterminística, ou seja, pelo fato do especialista ter poder de decisão na separação dos dados. Por fim, a tarefa de detecção de desvios estabelece normas que apontam anomalias, identificando fraudes em elementos que estão fora dos padrões ou fogem às regras.

<i>Tarefa</i>	<i>Descrição</i>	<i>Exemplos</i>
Agrupamento	Junta registros heterogêneos de dados em segmentos homogêneos, de acordo com similaridade entre eles.	Divisão de clientes de uma papelaria em grupos com hábitos de aquisição semelhantes.
Associação	Usadas para determinar quais itens tendem a co-ocorrerem (serem adquiridos juntos) em uma mesma transação.	Determinar quais os produtos costumam ser colocados juntos em um carrinho de supermercado.

Classificação	Constrói um modelo de algum tipo que possa ser aplicado a dados não classificados a fim de categorizá-los em classes.	Classificar pedidos de crédito Esclarecer pedidos de Seguros fraudulentos.
Correção de Desvios	Estabelece normas que apontam anomalias, identificando fraudes em elementos fora dos padrões.	Identificar o mês em que as notas dos alunos de determinada escola são mais baixas.

Tabela 1: Tarefas de Mineração de Dados

Considera-se técnicas de Mineração de Dados como sendo os fundamentos que irão permitir o desenvolvimento de algoritmos para procurar padrões escondidos nos dados. Cita-se como principais técnicas de Mineração: a Associação e a Segmentação/*Clustering*. A Associação contabiliza as ocorrências das combinações dos itens e depois quantifica a presença de um determinado item, representando as afinidades entre eles, sendo apresentada em formato texto (SCHNEIDER, 2002, p. 5). Exemplos de algoritmos que implementam regras de Associação: Apriori, Apriori Tid, AprioriHybrid, AIS, SETM e DHP (MIRANDA, 2004). Como na tarefa de Associação, uma variação da técnica de Associação são os Padrões Seqüenciais. Nesta técnica, sempre teremos informações relativas ao tempo da

transação, bem como à identificação do cliente que efetuou essa transação. Já a Segmentação/*Clustering* procura descobrir grupos homogêneos em uma base de dados, separando-a em segmentos de registros que compartilhem das mesmas propriedades (SCHNEIDER, 2002, p. 6). Outras técnicas são também conhecidas, como Árvores de Decisão – simples representação do conhecimento que classifica exemplos em um número finito de classes, usando CART e CHAID como algoritmos – , Indução de Regras – geração de um conjunto de condições não-hierárquicas, usando como algoritmos BIRCH e CLIQUE – e Redes Neurais – simulação computacional do sistema nervoso biológico, usando como algoritmos Perceptron, Rede BAM e Rede IAC, mas são usadas em menor escala (ÁVILA, PINTO, ZULLU JÚNIOR, ASSAD, 2003, p. 2).

Alguns *softwares* que realizam a mineração de dados são citados, com suas respectivas referências às técnicas implementadas por eles e algoritmos utilizados:

- a) *Intelligent Miner* da IBM: Utiliza-se das técnicas de Regras de Associação, Padrões Seqüenciais, Segmentação, Árvores de Decisão e Redes Neurais;
- b) *Enterprise Miner* – Tem-se as técnicas: Árvores de Decisão e Redes Neurais;
- c) Oracle 9i da Oracle, cujas técnicas são Regras de Associação e Redes Bayesianas;
- d) *DbMiner* - onde as técnicas utilizadas são: Regras de Associação e Árvores de Decisão.

5 Áreas de Aplicação e Evidências da Descoberta do Conhecimento

Em várias áreas e situações do mundo real já é freqüente a aplicação da

Descoberta do Conhecimento: de acordo com Moraes (2002, p. 35), evidencia-se a área de *Marketing* e Vendas na identificação de padrões de comportamento dos consumidores, inclusive por área demográfica e fidelidade; gerenciamento de relação com Clientes: Mineração comparativa de evolução de compras, produtos ou serviços oferecidos para definir política de preço altamente dinâmica. Na área financeira e de segurança, cita-se a identificação de padrões de fraude com cartões de crédito, planos de saúde, pagamento com impostos; identificação de características de correntistas e identificação da situação do mercado de ações. Ainda na esfera financeira, evidencia-se a mineração de séries temporais de valores na bolsa para especulação; mineração de associação e grupos de valores para definição de tipos de investimentos e planos de aposentadoria; previsão de inadimplência para definir política de empréstimo. No âmbito de combate a fraudes, destaca-se com as seguintes situações: mineração de séries excepcionais de pedidos de acesso a sites para segurança de sistemas, de transferência de fundos, para lavagem de dinheiro e sonegação de impostos; de chamadas para telecomunicações. No campo da Medicina e Biologia Molecular é citado o controle do comportamento dos pacientes, identificação de terapias de sucessos para diferentes tratamentos, comportamento dos usuários de planos de saúde e controle de medicamentos (fazendo previsão dos efeitos de um novo remédio composto a partir dos efeitos dos seus componentes em remédios testados). Na Segurança Pública, destaca-se a detecção de crimes: encontrando padrões nos dados que auxiliam na descoberta e prevenção de novos crimes. A situação muito usada para a Descoberta do Conhecimento é a Análise de Tendências,

onde, como ilustração, descreve-se o trabalho de empresas aéreas com o objetivo de aumentar cada vez mais a ocupação de seus vôos. Outro ponto de destaque da aplicação da Descoberta do Conhecimento é sobre a Análise de Mercado, onde se descobre produtos ou serviços que incentivam o consumo.

Com a realização de pesquisas na Internet para averiguação de outras situações de aplicação, lista-se outros enfoques como: o uso do Conhecimento na descoberta do comportamento do usuário da *web* (padrões ou regularidades que descrevem o seu perfil), cujas informações auxiliariam no projeto de remodelagem de sites, no gerenciamento dos recursos dos provedores de conteúdo na Internet, desenvolvimento de treinamentos e outros controles.

Em referência a outras pesquisas realizadas, destaca-se o projeto Oncoweb – Rede Piloto de Tele Saúde em Oncologia Pediátrica, liderado pela Escola Paulista de Medicina da Universidade Federal de São Paulo, onde o objetivo é a divulgação e troca de informações sobre Oncologia Pediátrica (UNIFESP, 2004). Neste contexto entra a parte da Descoberta do Conhecimento, onde as informações e as análises são formadas para melhoria técnico-profissional e divulgação aos hospitais parceiros.

Outra área que vem empregando a descoberta do conhecimento é a Bioinformática. Segundo Gomes, Lira, Sales, Silva, Ramalho Neto (2004), é uma ciência interdisciplinar que tem como objetivo desenvolver e aplicar técnicas computacionais nos estudos da genética, da biologia molecular e da bioquímica. Fundamentam a importância dos programas computacionais que trabalham na montagem do genoma a partir dos fragmentos de DNA enviados pelos

laboratórios de seqüenciamento. Além disso, são eles que permitem a comparação das seqüências descobertas com padrões genéticos já conhecidos e armazenados em bancos de dados, disponíveis na Internet. Existem grandes repositórios de dados genéticos e bioquímicos, como o GenBank, EMBL e NCBI, os quais contêm todas as seqüências de DNA conhecidas até hoje. Ainda na genética, outros exemplos de estudo podem ser citados: realização da mineração comparativa de seqüências de genes em células sãs e doentes e mineração de associações e grupos de genes baseado na sua co-ocorrência em várias classes de células. Na Universidade Federal de Alagoas, foi desenvolvido um projeto para identificar seqüências homólogas de genes relacionados com a resistência ao estresse hídrico da cana-de-açúcar, por meio de análises computacionais, efetuado no laboratório de Bioinformática/GEMPRO (Laboratório de Genética Molecular Genômica e Proteômica - Centro de Ciências Agrárias da Universidade Federal de Alagoas).

6 Utilização do Processo nas bases de dados do Exército Brasileiro

A proposta de contribuição deste artigo é a sugestão de que o Exército Brasileiro organize um projeto de Descoberta do Conhecimento para obter conhecimento e conseqüente análise estratégica sobre as bases de dados existentes, por parte do Alto Comando do Exército e dos órgãos de direção setorial.

O projeto, inicialmente, poderá realizar o processo *KDD* nas bases de dados administrativas. O Departamento Geral de Pessoal (DGP) contém as seguintes diretorias com as respectivas bases: Diretoria de Serviço Militar (DSM), Diretoria de Controle e Efetivo e

Movimentações (DCEM), Diretoria de Inativos e Pensionistas (DIP), Diretoria de Assistência Social (DAS) e Diretoria de Avaliação e Promoção (DAPROM). As bases da área de pagamento (SEF/CPEX) e de Dados de Cursos e Quadro de Cargos Previstos (QCD) que se encontram no Estado Maior do Exército (EME) também estão previstas, bem como as de Identificação Militar (IM) e o Módulo E1(E1). A Figura 3 mostra o esquema das bases de dados.

EXÉRCITO

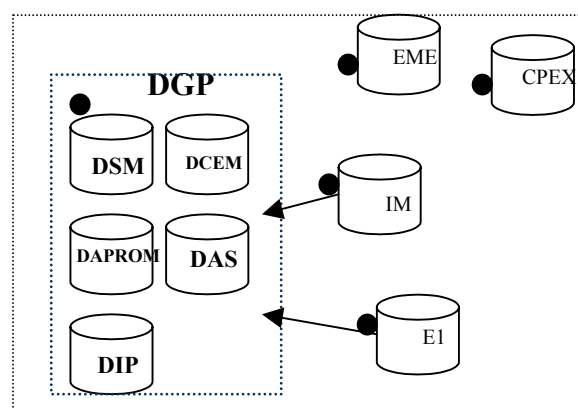


Figura 3: Esquema gráfico das bases de dados administrativas.

O projeto pode ser feito de duas formas:

- 1) Realizar o *KDD* em uma base única do Exército, integrando todas as bases mostradas na Figura 3. Este método apresenta a desvantagem de ter que aguardar todo o processo de integração, o que demandará muito tempo e, conseqüentemente, a utilização de conhecimento só será possível no final deste longo processo;
- 2) Realizar o *KDD*, isoladamente, em cada base de dados do DGP, EME, CPEX, IM e E1 (indicadas na Figura 3), obtendo o conhecimento imediato, sem precisar aguardar a integração. É importante mencionar que o conhecimento gerado nas bases IM e E1

poderá ser repassado para o DGP, visto que a informação é primordial para este.

Pode-se obter outros conhecimentos nas bases citadas acima como: No EME, é possível saber as vagas disponíveis nas Organizações Militares, de acordo com os postos/graduações. No CPEX é viável a descoberta do volume gasto nos últimos cinco anos com promoções e movimentações.

Na área de ensino, sugere-se também o trabalho de Descoberta do Conhecimento integrando as bases de dados dos Colégios militares, da mesma forma com as informações oriundas das Escolas de Formação. A Figura 4 mostra o esquema das referidas bases de dados: cada cilindro demonstra uma base de dados e, para generalizar, foi inserida a denominação XX, para que não fique difícil a compreensão. Os Colégios Militares citados foram os de Brasília-DF, Rio de Janeiro-RJ, Belo Horizonte- MG e Juiz de Fora-MG. As Escolas de formação referenciadas foram: a EsAEx (Escola de Administração do Exército), a AMAN (Academia Militar das Agulhas Negras), EsAO (Escola de Aperfeiçoamento de Oficiais) e EsSEX (Escola de Saúde do Exército).

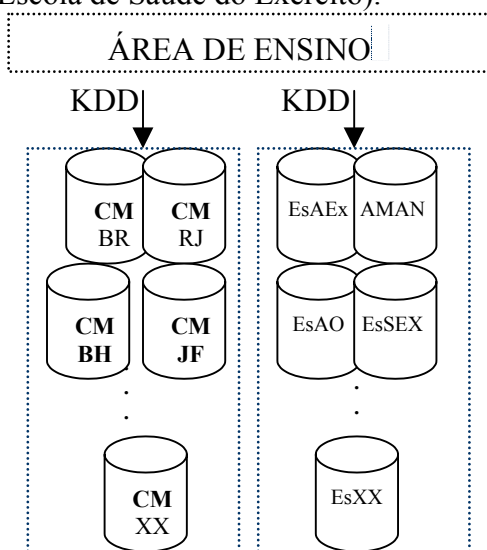


Figura 4: Esquema gráfico das bases de dados das áreas de ensino

Exemplo de alguns conhecimentos que podem ser explorados:

- a) Descobrir o melhor aluno, de acordo com as notas e conceitos, de determinados Colégios Militares e Escola de Formação;
- b) A menor nota ou faixa de alunos que obtiveram notas afins;
- c) A Escola que possui maior efetivo de alunos;
- d) Disciplinas afins entre os cursos oferecidos;
- e) Disciplina de maior rendimento;
- f) A percentagem de alunos, filhos de militares, dentre muitos outros.

Em todas as descobertas sugeridas há necessidade do envolvimento dos profissionais e a adesão do alto comando ou departamento/divisão/seção; presença dos especialistas do negócio e/ou do processo, os analistas de sistemas, os especialistas no *KDD* e Mineração de Dados. Estes profissionais e os usuários que divulgarão o conhecimento devem estar a par do cronograma e objetivos das etapas. Indica-se para boa administração e efetividade dos trabalhos um gerente de projeto, reportando diretamente à área estratégica ou à Secretaria de Tecnologia da Informação.

7 Considerações Finais

Este trabalho procurou valorizar e mostrar de forma estratégica a utilização da Descoberta do Conhecimento em bases de dados do Exército Brasileiro, através do processo denominado *KDD* (*Knowledge Discovery in Databases*). Assim, poder-se-á haver maior produtividade nos planejamentos e, principalmente, na tomada de decisões cabíveis de forma precisa e correta em cada situação. Recomenda-se, como subsequente a este

artigo, o estudo, a avaliação de ferramentas utilizadas na quinta etapa do processo, que é o *Data Mining* ou Mineração de Dados, bem como estudos aprofundados baseados nas propostas sugeridas na seção 6.

Como fatores essenciais para o sucesso do projeto, descreve-se baixa percentagem de erros nos dados e – primordial para garantia do anterior – o comprometimento de toda a Instituição/Organização, já que, somente com profissionalismo e efetividade dos recursos humanos envolvidos, o projeto irá se concretizar e atender as expectativas. Caso seja uma nova metodologia de trabalho, o treinamento e os estudos de casos de sucesso também são passos fundamentais para garantia dos objetivos. Os casos de sucesso ajudarão na avaliação e customização do projeto a ser desenvolvido.

Os resultados obtidos com a ferramenta são: melhores lucros, menores custos, aproveitamento do tempo, melhoria da qualidade e do processo da tomada de decisão estratégica.

Dispondo-se da informação, do conhecimento no momento certo, uma Instituição ou empresa estará pronta para as mudanças, crescimentos, desafios e competitividade.

Referências

ÁVILA, Ana Maria Heuminski de; PINTO, Hilton Silveira; ZULLU JÚNIOR, Jurandir; ASSAD, Eduardo Delgado. **Estimativa de Precipitação Através de Imagens do Satélite GOES-8 Utilizando Redes Neurais**. Universidade Estadual de Campinas, UNICAMP, Campinas, 2003. Disponível em http://lagavulin.ltid.inpe.br:1905/col/ltid.inpe.br/sbsr/2003/03.27.11.57/doc/11_302.PDF> Acesso em: 29 mar. 2004.

BARIONI, Maria Camila Nardini. **Visualização de Operações de Junção em Sistemas de Bases de dados para Mineração de Dados**. Instituto de Ciências Matemáticas e de Computação, USP, São Carlos, 2002. Disponível em <http://www.teses.usp.br>> Acesso em: 29 mar. 2004.

COME, Gilberto de. **Contribuição ao Estudo da Implementação de Data Warehousing: Um Caso no Setor de Telecomunicações**. Faculdade de Economia, Administração e Contabilidade, Departamento de Administração, USP, São Paulo, 2001. Disponível em <http://www.teses.usp.br>> Acesso em: 29 mar. 2004.

FIGUEIRA, Rafael. **Mineração de Dados e Banco de Dados Orientados a Objetos**. COPPE Sistemas, UFRJ, Rio de Janeiro, 1998. Disponível em <http://www.cos.ufrj.br/~rafael/mestrado/bdnc/monografia.pdf>. Acesso em: 29 mar.2004.

GOMES, Suelly da Silva; LIRA, Giselle Amorim; SALES, Janice Silva; SILVA, Denise Wanderlei; RAMALHO NETO, Eduardo. **Pesquisa de Genes da Cana-de-Açúcar Relacionados ao Estresse Hídrico, Utilizando Ferramentas de Bioinformática**. I Jornada Científica de Farmácia, UFAL. Disponível em http://www.propep.ufal.br/Eventos/I%20Jornada%20Cientifica%20de%20Farmacia/Resumo_25.htm>. Acesso em: 21 abr. 2004.

MIRANDA, Dhalila. **Algoritmo Apriori**. Disponível em <http://inf.aedb.br/datamining/paginas/apriori.htm>>. Acesso em:18 mai. 2004.

MORAES, Sérgio. **Descoberta do Conhecimento em Base de Dados. uma breve visita**. UCB, Brasília, 2002. Disponível em <http://www.fazenda.gov.br/ucp/pnafe/doc/s/p2-S%C3%A9rgio.pps>> . Acesso em: 05 abr. 2004.

OLIVEIRA, Robson Butacca Taborelli. **O Processo de Extração do**

Conhecimento de Base de Dados Apoiado por Agentes de Software. Instituto de Ciências Matemáticas e de Computação, USP, São Carlos, 2000. Disponível em <http://www.teses.usp.br>. Acesso em: 05 abr. 2004.

SANTOS, Manuel Felipe. **Descoberta de Conhecimento em Base de Dados.** Departamento de Sistemas de Informação, Escola de Engenharia, Universidade do Minho, Coimbra-Portugal, 2002. Disponível em http://piano.dsi.uminho.pt/disciplinas/M_SIAID/dm.PDF. Acesso em: 05 abr. 2004.

SCHNEIDER, Luís Felipe. **Mineração de Dados: - Conceitos.** Departamento de Agronomia, UFRS, Porto Alegre, 2002. Disponível em http://www.inf.ufrgs.br/~clesio/cmp151/cmp15120021/artigo_lfelipe.pdf. Acesso em: 05 abr. 2004.

SILVA, Edilberto Magalhães Silva. **Avaliação do Estado da Arte e Produtos Data Mining.** Mestrado em Informática, UCB, 2000. Disponível em http://www.edilms.hpg.ig.com.br/download/Trabalho_TSI_Edilberto_Data_Minig.pdf. Acesso em: 05 abr. 2004.

UNIFESP, Universidade Federal de São Paulo, Escola Paulista de Medicina, Setor de Telemedicina. **Projeto ONCOWEB.** Disponível em <http://www.unifesp.br/dis/set/project/onco/>. Acesso em: 06 mai. 2004.