

**LIGA DE ENSINO DO RIO GRANDE DO NORTE
FACULDADE NATALENSE PARA O DESENVOLVIMENTO
DO RIO GRANDE DO NORTE**

**APLICAÇÃO DA FERRAMENTA DATA MINING EM ORGANIZAÇÕES DE
MÉDIO PORTE: UM ESTUDO DE CASO NA FARN.**

**APLICAÇÃO DA FERRAMENTA DATA MINING EM ORGANIZAÇÕES DE
MÉDIO PORTE: UM ESTUDO DE CASO NA FARN.**

Monografia apresentada ao Curso de Graduação em Bacharelado em Sistema de Informação da Faculdade Natalense para o Desenvolvimento do Rio Grande do Norte como requisito parcial para a obtenção do grau de Bacharel em Sistema de Informação.

Aylzemara Costa e Silva

NATAL

2002

AYLZEMARA COSTA E SILVA

LIGA DE DESENVOLVIMENTO DO NORTE
FACULDADE NATALENSE PARA O DESENVOLVIMENTO
DO RIO GRANDE DO NORTE

APLICAÇÃO DA FERRAMENTA DATA MINING EM ORGANIZAÇÕES DE
MÉDIO PORTE: UM ESTUDO DE CASO NA FARN.

**APLICAÇÃO DA FERRAMENTA DATA MINING EM ORGANIZAÇÕES DE
MÉDIO PORTE: UM ESTUDO DE CASO NA FARN.**

Aylzemara Costa e Silva.

Monografia apresentada e aprovada em _____ de _____ de _____
Pela banca examinadora composta pelos seguintes membros:

Monografia apresentada ao Curso de
Graduação em Bacharelado em Sistema
de Informação da Faculdade Natalense
para o Desenvolvimento do Rio Grande
do Norte como requisito parcial para a
obtenção do grau de Bacharel em Sistema
de informação

NATAL

2002

RB-13715

FARN

BIBL. TECA

Nº Reg. M000223

04/05/05

SILVA, Aylzemara Costa e

Aplicação da ferramenta Data Mining em Organizações de médio porte: um estudo de caso na FARN
Aylzemara Costa e Silva – Natal: FARN, 2002.

35 p.

Monografia (Bacharelado em Sistema de Informação) – Faculdade Natalense para o Desenvolvimento
do Rio Grande do Norte..

1. Ferramenta para desenvolvimento de programas – monografia. 2 – Banco de Dados – monografia. I
– Título.

RN/ FARN/ BR

CDU.004.42

M

004.4'2

S578a

2002

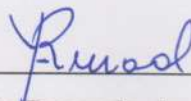
**LIGA DE ENSINO DO RIO GRANDE DO NORTE
FACULDADE NATALENSE PARA O DESENVOLVIMENTO
DO RIO GRANDE DO NORTE**

**APLICAÇÃO DA FERRAMENTA DATA MINING EM ORGANIZAÇÕES DE
MÉDIO PORTE: UM ESTUDO DE CASO NA FARN.**

Aylzemara Costa e Silva.

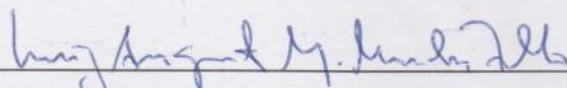
Monografia apresentada e aprovada em ____ de ____ de ____.

Pela banca examinadora composta pelos seguintes membros.

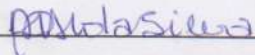


Ytalo Rosendo do Amaral

Orientador



Luiz Augusto machado de Mendes Filho



Livia Maria Martins da Silva

Natal, ____ de ____ de ____.

A DEDICATÓRIA.

Agradeço por meio deste trabalho a Deus em primeiro lugar, pela oportunidade e pela capacidade que tive para a realização do mesmo. A meus pais por toda a dedicação e obstáculos enfrentados ao longo desses quatro anos, sempre ao meu lado me incentivando e até mesmo ajudando-me a seguir adiante quando alguma coisa parecia não conseguiria.

Também, agradeço a meu namorado Leonardo, pela compreensão, paciência e apoio durante as horas mais difíceis, fazendo com que eu me concentrasse no meu trabalho e conseguindo concluir este projeto.

Ainda, agradeço a todos os meus professores, pelos ensinamentos ministrados e pelo estímulo aos meus estudos e pesquisas.

A minha Coordenadora Lívia, pela orientação e apoio durante todo o processo de elaboração e conclusão deste trabalho. Aos meus pais dedico este trabalho, como resultado de meu esforço e também do apoio, amor e compreensão que deles recebi em todos os projetos e obstáculos da minha vida.

Ao meu orientador, em primeiro lugar, agradeço por toda a orientação e apoio durante todo o processo de elaboração e conclusão deste trabalho. Aos meus pais dedico este trabalho, como resultado de meu esforço e também do apoio, amor e compreensão que deles recebi em todos os projetos e obstáculos da minha vida.

Aos meus amigos e familiares, agradeço por toda a compreensão e apoio durante todo o processo de elaboração e conclusão deste trabalho. Aos meus pais dedico este trabalho, como resultado de meu esforço e também do apoio, amor e compreensão que deles recebi em todos os projetos e obstáculos da minha vida.

A Faculdade Natalense de Rio Grande do Norte, particularmente aos professores, funcionários e demais colaboradores, pela atenção e apoio que me ofereceram para a realização e conclusão deste trabalho acadêmico e científico, que é a produção de conhecimento.

E a todos que, de alguma forma, contribuíram na elaboração deste trabalho.

AGRADECIMENTOS.

Agradeço por meio desse trabalho a Deus em primeiro lugar, pela oportunidade e pela capacidade que tive para a realização do mesmo. A meus pais por toda a dedicação e obstáculos enfrentados ao longo desses quatro anos, estando sempre ao meu lado me incentivado, e até mesmo ajudando-me a seguir enfrente que imaginei em alguns instantes que não conseguiria.

Também, agradeço a meu namorado Leonardo, pela compreensão e paciência depositada em mim nas horas mais difíceis, fazendo com que eu me enchesse de entusiasmo resultando na conclusão desse projeto.

Assim, agradeço a todos os meus professores, pelos ensinamentos ministrados e pelo estímulo aos desafios profissionais e acadêmicos que deles recebi.

A minha Coordenadora Livia Maria Martins da Silva, agradeço pela disponibilidade, dedicação e amizade em mim depositada ao longo desses anos acadêmicos.

Ao meu orientador, em especial o Prof. Ytalo Rosendo do Amaral, agradeço pela paciência e dedicação depositada em mim ao longo de toda a realização desse trabalho, pelos pulxões de orelhas dados em algumas horas que mereci, que me serviram de incentivo para continuar e chegar a conclusão do meu trabalho acadêmico. Pelas críticas feitas no trabalho ao longo de toda a sua realização, sabendo perceber os erros e acertos deste trabalho, fazendo sempre o melhor, quando possível.

Aos meus amigos e companheiros, conquistados ao longo desses quatro anos, que ao fim de todo esse obstáculo que passamos ficará a saudade dos bons momentos passados juntos. Aos meus amigos também conquistados ao longo do meu estágio no TRE, agradeço pelos ensinamentos e pela oportunidade de colocar todo o meu conhecimento acadêmico em pratica, sem falar em todo o conhecimento que lá adquiere.

À Faculdade Natalense do Rio Grande do Norte, particularmente nas pessoas de seus Diretores, Professores e Funcionários, pela presteza e suporte que me ofereceram para a realização e conclusão desta oportunidade produtiva e valorosa que é a produção de conhecimento.

E a todos que, de alguma forma, contribuíram na elaboração deste trabalho.

LISTA DE CONTEÚDOS SUMÁRIO

1. INTRODUCAO.....	10
1.1- OBJETIVOS.....	13
1.2- JUSTIFICATIVA.....	13
2. REFERENCIAL TEORICO.....	14
2.1- A ERA DA INFORMAÇÃO.....	14
2.2- DEFINIÇÃO.....	14
2.3- DATA MINING.....	15
2.3.1- Objetivos do Data Mining.....	16
2.3.2 – Funcionamento do Data Mining.....	16
2.3.3 – Áreas de Aplicações do Data Mining.....	18
2.3.4- Técnicas de Data Mining.....	20
2.3.5 - O Processo de Implementação do Data Mining.....	21
2.3.6 – Algumas Barreiras ao uso do Data Mining.....	23
3. ANALISE DOS RESULTADOS.....	25
3.1- WIZRULE.....	26
3.2 – SUPERQUERY.....	27
4. CONCLUSAO.....	30
4.1- SUGESTÃO.....	31
5. REFERENCIA BIBLIOGRAFICA	32
6. BIBLIOGRAFIA CONSULTADA.....	34

LISTAS DE ILUSTRAÇÕES

Figura1- Etapas para realização da mineração de dados.....	13
Gafico1- Representação gráfica dos resultados encontrados com a mineração de dados referentes as tabelas: Empréstimo X Curso.....	29

LISTAS DE TABELAS E QUADROS

Tabela 1- Valores estatísticos para material contido na biblioteca da FARN.....	29
Tabela 2- Valores estatísticos das edições contidas no acervo da biblioteca da FARN.....	28
Tabela 3- Requisitos mínimos para a utilização das ferramentas Wizrule e SuperQuery.....	26
Tabela 4 - Informações sobre custo e disponibilidade das ferramentas Wizrule e SuperQuery.....	30

RESUMO.

O processo de descoberta do conhecimento e mineração de dados vem sendo amplamente difundidos nas diversas áreas com o objetivo de encontrar informações úteis que normalmente não estão visíveis nos gigantescos bancos de dados. O Data Mining prevê tendências de comportamento de clientes, através de informações contidas nas bases de dados das organizações. Este trabalho tem o objetivo de apresentar, de uma forma clara, os principais conceitos e técnicas usadas na tecnologia Data Mining que é composta por um conjunto de ferramentas. A análise foi realizada na mineração dos dados no banco de dados da biblioteca da FARN, utilizando-se dos seguintes softwares: Wizrule e SuperQuery para a obtenção de resultados e análise nas tendências dos usuários pertencentes ao banco de dados.

Palavras –chaves

Mineração de dados, banco de dados, organizações, tecnologia.

Knowledge and Data Mining's discovery process come being widely broadcast on diverser a areas intending to find useful informations that normally are not visible in mammoth data-bank.. Data Mining predicts customers tendencies through information contained in organizations's data-bank. This work intends to introduce, by a clar way, principal concepts and technigues used on Data Mining technology that is formed by a tools set. Analysis was realized in FARN library data-bank data mining using the following softwares: Wizrule and SuperQuery to obtain results and to analyse users tendency in data-bank.

KeyWords

Data Mining, Data-bank, organizations, tecnology.

Surgindo assim, a inteligência competitiva, que é uma atividade de gestão estratégica de informação que tem como objetivo permitir que os tomadores de decisão se antecipem sobre as tendências dos mercados e a evolução da concorrência, detectando e analisando as ameaças e oportunidades que se apresentam no desenvolvimento da organização.

Pois as informações de um sistema de informações visam principalmente tornar a organização mais ágil, capaz de definir com mais clareza a missão da organização no mercado, planejar estratégias, pôr em prática e avaliar os resultados, monitorar os movimentos dos concorrentes, definir os comandantes, reestruturação de departamentos e mercados, etc.

Uma outra área bastante conhecida é que está relacionada a gestão empresarial é o Business Intelligence. Que tem origem no conhecimento do negócio na era da computação global e das comunicações on-line, sendo chamado por dois de hoje de Business Intelligence (BI) ou Inteligência de Negócios.

Segundo TYSON apud ANDRINI (2002), Inteligência de Negócios é

1. INTRODUÇÃO.

A cada ano, companhias acumulam mais e mais informações em seus bancos de dados. Como consequência, estes bancos de dados passam a conter verdadeiros tesouros de informação sobre vários procedimentos dessas companhias. Toda essa informação pode ser usada para melhorar seus procedimentos, permitindo que a organização detecte tendências e características disfarçadas, e reaja rapidamente a um evento que ainda pode estar por vir. Pois praticamente todas as organizações de médio e grande porte estão utilizando sistemas informatizados para realizar seus processos mais importantes, o que com o passar do tempo acaba gerando uma enorme quantidade de dados relacionados aos negócios, mas não relacionados entre si.

Contudo, a simples armazenagem e recuperação dessa informação já traz um grande benefício, pois agora já não é mais necessário procurar informação em volumosos e ineficazes arquivos de papel, pois com o avanço da tecnologia o armazenamento é feito em poderosos e gigantescos bancos de dados. Mas apesar de toda essa importância no armazenamento da informação a quantidade de informação que já vem sendo armazenada nos banco de dados existentes está explodindo, e ultrapassa a habilidade técnica e a capacidade humana na sua interpretação. Banco de dados são medidos hoje em gigabytes e terabytes. A necessidade de transformar estes terabytes de dados em informações é de extrema importância para os administradores das organizações.

Surgindo assim, a inteligência competitiva, que é uma atividade de gestão estratégica da informação que tem como objetivo permitir que os tomadores de decisão se antecipem sobre as tendências dos mercados e a evolução da concorrência, detectando e avaliando as ameaças e oportunidades que se apresentam no desenvolvimento da organização.

Pois as informações de um sistema de informações visam principalmente tornar a organização mais ágil, capaz de definir com mais clareza a situação da organização no mercado, planejar estratégias, pontos fortes, e fracos, recursos, movimentos dos competidores, definição dos consumidores, mensuração de demanda e mercados, etc.

Uma outra área bastante conhecida e que estar relacionada a gestão empresarial é o Business Intelligence. Que teve origem no conhecimento do negócio na era da competição global e das comunicações on-line, sendo chamado nos dias de hoje de Business Intelligence (BI) Ou Inteligência de Negócios.

Segundo TYSON apud ANDRINI (2002), Inteligência de Negócios é:

SILVA, Aylzema Costa e

Aplicação da ferramenta Data Mining em Organizações de médio porte: um estudo de caso na FARN/
Aylzema Costa e Silva – Natal: FARN, 2002.

35 p.

Monografia (Bacharelado em Sistema de Informação) – Faculdade Natalense para o Desenvolvimento
do Rio Grande do Norte.

1. Ferramenta para desenvolvimento de programas – monografia. 2 – Banco de Dados – monografia. I
– Título.

RN/ FARN/ BR

CDU.004.4'2

“Um processo que envolve a coleta, análise e validação de informações sobre concorrentes, clientes, fornecedores, candidatos potenciais à aquisição, candidatos à joint-venture e alianças estratégicas. Inclui também eventos econômicos, reguladores e políticos que tenham impacto sobre os negócios da empresa. O processo de Inteligência de Negócios analisa e valida todas essas informações e as transforma em conhecimento estratégico”

Segundo o Portal BR-Business (2002), os benefícios oferecidos pelo Business

Intelligence são os de:

- Antecipar mudanças no mercado;
- Antecipar ações dos competidores;
- Descobrir novos ou potenciais competidores;
- Aprender com os sucessos e as falhas dos outros;
- Conhecer melhor suas possíveis aquisições ou parceiros;
- Conhecer novas tecnologias, produtos ou processos que tenham impacto no seu negócio;
- Entrar em novos negócios;
- Rever suas próprias práticas de negócio;

Algumas das ferramentas que fazem parte de Business Intelligence:

- Planilha eletrônica
- Geradores de queries e de relatórios
- Sistemas de suporte às decisões (DSS)
- Sistemas de informação executiva (EIS)
- Data Warehouse
- Data Mart
- On-line análise (OLAP)
- Data Mining
- Suites corporativas de BI

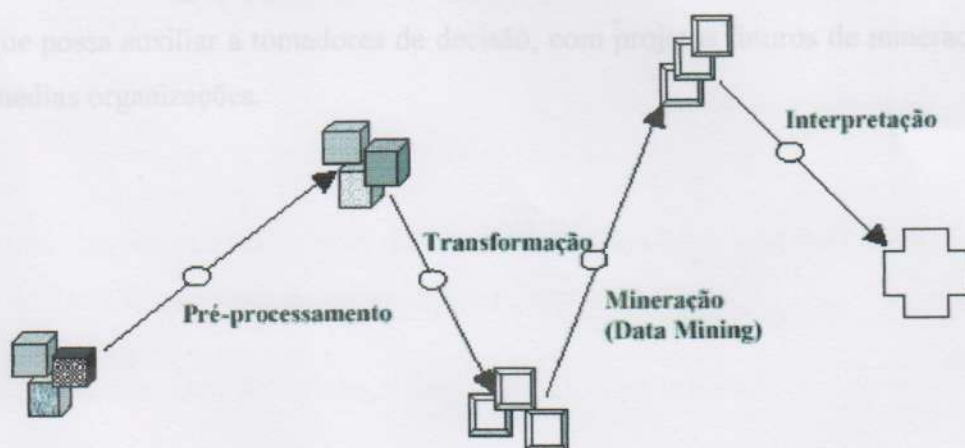
Desta forma a tecnologia da informação estará exercendo seu grande papel, que é o de fornecer informações de qualidade, e deixando de ser um simples amontoado de dados.

Sabe-se que uma das armas mais poderosas e competitivas no mercado é a capacidade de conhecer e oferecer produtos adequados às necessidades individuais dos clientes. E quando se descobre isso, pode-se estabelecer estratégias para otimizar os resultados financeiros da organização. Contudo, embora essa idéia básica seja facilmente compreensível, fica sempre uma dúvida sobre como um sistema é capaz de obter esse tipo de relação que para as organizações é essencial para que a mesma permaneça no mercado será preciso que sempre

elas tenham a capacidade de competir uma com as outras. Sendo um dos diferenciais mais importantes há obtenção da informação no momento certo para que possa tomar a melhor decisão no momento exato, fazendo com que a organização possa obter bons resultados em seus negócios. Para as organizações a importância de cada decisão tomada é a permanência da mesma no mercado de igual para igual com os seus concorrentes, pois a decisão mal tomada pode colocar a vida útil da organização em jogo.

Sabendo de tudo isso surge o Data Warehouse, ou em português, Armazém de Dados, que pode ser definido como um banco de dados especializado, o qual integra e gerencia o fluxo de informações a partir dos bancos de dados corporativos e fontes de dados externos à organização. Um Data Warehouse é construído para que tais dados possam ser armazenados e acessados de forma que não sejam limitados por tabelas e linhas estritamente relacionais. A função do Data Warehouse é tornar as informações corporativas acessíveis para o seu entendimento, gerenciamento e uso. Como o Data Warehouse separado dos bancos de dados operacionais, as consultas dos usuários não impactam nestes sistemas, que ficam resguardados de alterações indevidas ou perdas de dados. O Data Warehouse não é como um software, que pode ser comprado e instalado em todos os computadores da organização em algumas horas, na realidade sua implantação exige a integração de vários produtos e processos. (DAL'ALBA, 2002).

Figura 1- Etapas para realização da mineração de dados.



2. REFERENCIAL TEÓRICO

1.1 OBJETIVOS

Este trabalho tem o objetivo de apresentar, de uma forma clara e objetiva, os principais conceitos e técnicas usadas na tecnologia Data Mining que é composta por um conjunto de ferramentas que estão sendo amplamente difundida e aperfeiçoada. Essa tecnologia tem como objetivo secundário de analisar e prever tendências de comportamento de clientes, através de informações contidas nas bases de dados das organizações.

Um outro objetivo considerado secundário é de se fazer um estudo das técnicas e ferramentas existentes de Data Mining, fazendo um estudo das ferramentas que podem ser utilizadas com sucesso em pequenas e medias organizações para que com isso se possa analisar se é o não favorável a utilização essas ferramentas de Data Mining. Para isso foi escolhido um banco de dados da própria instituição (FARN), para que a mineração de dados fosse realizada e com isso identificado as vantagens e desvantagens da mineração de dados nas pequenas e medias organizações.

1.2- JUSTIFICATIVA

O que me despertou para elaboração desse trabalho, foi a necessidade de se fazer um levantamento e descrição das técnicas existentes de Data Mining no mercado e o estudo de ferramentas que tenha como prevê e identificar, se é ou não, possível a mineração de dados nas pequenas e medias organizações. Tendo a justificativa final de se levantar um trabalho acadêmico que possa auxiliar a tomadores de decisão, com projetos futuros de mineração em pequenas e medias organizações.

2.2- DEFINIÇÕES

Para entendermos os fundamentos das ferramentas Data Mining, é importante entender os conceitos de Data Warehouse, Banco de Dados e Data Mart.

Segundo Orin do Dall'olio e Robinson Grips (2007), Data Warehouse:

"É um banco de dados contendo dados extraídos do sistema de produção da empresa, foram selecionados e depurados, para a programação de transações. Em geral, um Data Warehouse engloba a consolidação de outros recursos de dados que não se atualizam no

2. REFERENCIAL TEORICO

2.1- A ERA DA INFORMAÇÃO:

Ao longo da história da humanidade, a informação (e conseqüentemente o conhecimento) tem sido o fator decisivo em todos os confrontos e disputas em que o homem se envolveu. Pois a informação desempenha um papel fundamental em todos os setores da sociedade. O desenvolvimento e sucesso das organizações baseiam-se, cada vez mais, na capacidade de coletar, tratar, interpretar e utilizar a informação de forma eficaz.

Segundo Santos (1999), "para o homem das cavernas, a sobrevivência implicou o conhecimento (e domínio) do fogo. Para o império Romano, o conhecimento tecnológico (como a construção de estradas, pontes e aquedutos), permitiu gerir (a nível militar e econômico) durante vários séculos um enorme império. Na época dos descobrimentos, a informação de anteriores expedições foi fundamental para novos descobrimentos".

Hoje em dia vivemos numa época altamente consumista e as pessoas querem consumir cada vez mais informação, pois percebem a sua importância em nosso dia-a-dia. Mais com o rápido evoluir da tecnologia, também veio aumentar esta tendência, pois com toda essa evolução tornou-se cada vez mais acessível o acesso à informação desejada. Pois hoje em dia a informação, consiste na transmissão rápida de dados para os mais diversos lugares do mundo e para o maior número possível de pessoas e organizações. Um claro exemplo disto é a própria internet. Com isso pode dizer que qualquer organização que queira sobreviver, não só não pode ignorar informação importante que disponha, como tem de procurar meios de conseguir mais informação.

2.2- DEFINIÇÕES:

Para entendermos os fundamentos das ferramentas Data Mining, é importante citarmos os conceitos de Data Warehouse, Bases de Dados e Data Mart.

Segundo Osvaldo Dalfofo e Robson Gripa (2002), Data Warehouse:

“É um banco de dados contendo dados extraídos do ambiente de produção da empresa, foram selecionados e depurados, além de otimizados para o processamento de transações. Em geral, um Data Warehouse requer a consolidação de outros recursos de dados que não os armazenados nos

bancos de dados relacionais, incluindo informações provenientes de planilhas eletrônicas, documentos textuais, etc.”

Segundo Torres (2002),

Bases de Dados é o responsável pela gestão da informação. Quando o programador cria as aplicações com que o utilizador final irá trabalhar para inserir, consultar ou apagar a informação estas comunicarão com o Sistema de Gestão à Bases de Dados para realizar essas tarefas. Assim haverá independência entre as aplicações e a estrutura de dados.

Segundo Assunção (2002) Data Mart é:

“É um subset do Data Warehouse. Construído com um escopo menor de informações segmentado para uma determinada área ou assunto. Por ser menor, possibilita a Análise Multidimensional, com os cruzamentos e visões previamente calculados, visando aumentar a velocidade na consulta das informações.”

2.3- DATA MINING:

De acordo com Zanasi (2002),

"Data Mining são todas as técnicas que permitem "extrair" conhecimento que, de outra maneira, permaneceria escondido nas grandes bases de dados. Os algoritmos computacionais que tornam isso possível, estão baseados na matemática e em novas e inteligentes maneiras de fazer uso do poder de processamento do computador. Para a exposição simples dessas técnicas e de suas aplicações".

Segundo Brusso (2002), Data Mining ou Mineração de Dados é :

“É uma das etapas do processo de descoberta de conhecimento em bancos de dados e têm por objetivo extrair informações implícitas e potencialmente úteis de dados . Normalmente todo este processo consiste em uma série de etapas, que inicia com a seleção do conjunto ou amostra dos dados a serem processados. Estes dados podem ser submetidos a uma etapa de pré-processamento onde são transformados para um formato adequado para o algoritmo de mineração, que procura por regras ou padrões ocultos nos dados.”

Uma técnica muito difundida nos dias de hoje é a mineração de dados ou Data Mining que consiste basicamente na extração conhecimento interessante e útil para tomada de decisão. O Data Mining não requer que o usuário 'pilote' a ferramenta ao longo do processo de análise dos dados ou que administre o processo ao longo de seu andamento. Fornecidos os dados adequados no início do processo, o Data Mining pode fazer sentido a grupos de dados achando tendências ou padrões 'escondidos' nos mesmos, e apresentá-los ao usuário num formato compreensível.

2.3.1- Objetivos do Data Mining:

- **Explanatório:** Explicar algum evento ou medida observada, tal como porque do aumento nas vendas de cervejas e fraldas descartáveis nos fins de semana; AURÉLIO (2002).
- **Confirmatório:** confirmar uma hipótese. Uma companhia de seguros, por exemplo, pode querer examinar os registros de seus clientes para determinar se família de duas rendas tem mais probabilidade de adquirir um plano de saúde do que famílias de uma renda; AURÉLIO (2002).
- **Exploratório:** analisar os dados buscando relacionamentos novos e não previstos. Uma companhia de cartão de crédito pode analisar seus registros históricos para determinar que fatores estão associados a pessoas que representam risco para créditos. AURÉLIO (2002).

2.3.2 - Funcionamento do Data Mining .

O interesse pela mineração de dados ou Data Mining surgiu devido ao grande armazenamento de informações em gigantescas bases de dados da grande necessidade no aproveitamento dessas informações nas chamadas tomadas decisão. Um fator que contribuiu e muito para o aumento do interesse em mineração de dados foi o desenvolvimento das técnicas para o auxílio à tomada de decisão - são elas: as redes neurais artificiais, algoritmos genéticos, lógica indutiva, métodos estatísticos, análise de cluster, regressão, árvore de decisão e entre outras, que auxiliam em todo o processo de mineração dos dados. Pois na fase de mineração de dados o executor da tarefa pode utilizar várias ferramentas e técnicas para

que o seu objetivo seja bem sucedido. Uma das ferramentas mais utilizadas são aquelas baseadas em consulta a bases de dados, linguagem SQL, para a obtenção de uma análise preliminar dos dados armazenados em todo o processo da construção do Data Warehouse.

O Processo de mineração de dados se dar desde da extração dos dados nas bases de dados, passando por uma limpeza e transformação dos dados. Esta fase tem o objetivo da retira de informações desnecessárias do banco de dados, sem que isso altere nenhuma das informações já armazenadas, Pois a limpeza dos dados significa não deixar que dados incorretos alterem ou prejudiquem todo o andamento da mineração dos dados.

Segundo HARRISON (1998),

“Não existe uma técnica de Data Mining que resolva todos os problemas. A familiaridade com as técnicas é necessária para proporcionar a melhor abordagem de acordo com os problemas apresentados. Portanto, para cada classe de aplicação deve-se aplicar um conjunto de algoritmos com o objetivo de extrair padrões e relações dentro de uma base de dados. Entre as principais famílias de algoritmos de Data Mining, destacamos regras de associação, classificação ou regressão, agrupamento ("clustering") e outras”

Vejamos exemplos de cada:

▪ Associação

As associações são relacionamentos entre itens de dados armazenados. O objetivo da operação é encontrar tendências, a partir de um grande número de transações, que possam ser usadas para entender e explorar padrões de comportamentos dos dados.

– **Exemplo** (típico): transações de cartões de crédito.

• Classificação ou regressão

A classificação ou regressão usa dados existentes para criar modelos de comportamentos de variáveis, caracterizando-se por ser uma técnica que basea-se na aplicação de um conjunto de exemplos pré-classificados para a elaboração de um modelo capaz de classificar uma população maior de registros. Podemos citar com exemplo de análise a detecção de fraudes e aplicações de risco. Os algoritmos de classificação incluem árvores de decisão ou redes neurais.

– **Exemplos**: pedidos de crédito, pedidos de indenização fraudulentos.

▪ Agrupamento (Clustering)

O Agrupamento em clusters envolve segmentar a informação disponível em conjuntos definidos e homogêneos baseando-se em atributos específicos. O conceito de clustering já tem uma longa história em estatística, mas o que tem de novo em Data Mining é o fato de poder também ser aplicada a itens não numéricos. Os resultados de uma operação de clusterização podem ser usados de duas diferentes maneiras: para produzir um sumário da base de dados ou como dados de entrada para outras técnicas; por exemplo, classificação, pois um cluster é um grupo menor e de manuseio mais fácil por parte de algoritmos de classificação.

– **Exemplo:** segmentação de mercados, descoberta de sub-grupos, de nichos de mercado, organização de documentos.

2.3.3 – Áreas de Aplicações do Data Mining :

a) Vendas (Marketing):

- a.1) Com análise de compra de seus clientes através dos cartões de crédito;
- a.2) Identificar padrões de comportamento dos consumidores;
- a.3) Prever quais consumidores serão atingidos nas campanhas de marketing;

– **Exemplo:** A empresa Express analisa os padrões de compra dos seus clientes, possuidores de cartões de crédito, e usa informação para oferecer promoções dirigidas individualmente. (SANTOS, 2002).

Como a American Express temos em Portugal a SONAE com o produto Cartão Universo que permite guardar as referências de todas as compras efetuadas pelos respectivos clientes nos hipermercados CONTINENTE e nas lojas Modelo. Bem como o Petrogal com o cartão Fast GALP, controlando, ou melhor, guardando a informação das transações efetuadas pelos clientes nas lojas de combustíveis e outros produtos

b) Seguros e Planos de Saúde:

- b.1) Determinar quais procedimentos médicos são requisitados ao mesmo tempo;
- b.2) Prever quais consumidores comprarão novas apólices;
- b.3) Identificar comportamentos fraudulentos;

– **Exemplo:** O Health-Kefir é um excelente exemplo de como se podem reduzir os custos associados aos cuidados de saúde e simultaneamente melhorar esses mesmos cuidados,

parece contraditório, mas não; vejamos o funcionamento do KEFIR: o sistema apenas seleciona as condições interessantes como por exemplo as mulheres com uma gravidez normal não são selecionadas, enquanto que as que tiverem nascimentos prematuros são. Assim podem ser canalizados os esforços para intervenções médicas (de rotina) que promovam a redução da taxa de nascimentos prematuros. Evitando as despesas maiores dos cuidados excepcionais necessários para os prematuros. (SANTOS, 2002).

c) Finanças:

c.1) Determinar no suporte à decisão na área dos empréstimos bancários no intuito de prevenir a fraude;

c.2) Determinar que clientes teriam condições de fazer um empréstimo;

– **Exemplo:** O BBP-Banco Privado Português sentiu a necessidade de um suporte à decisão de avaliação de perfis de riscos para os investimentos financeiros dos seus clientes. Assim contratou a equipe da Novabase Suporte à Decisão para obter uma solução. A Novabase Suporte à Decisão sugeriu, para esse problema, a utilização de técnicas de Data Mining para a criação dos modelos de identificação de estratégias de investimentos. Esse processo permitia ao BPP a utilização da sua experiência com os seus clientes e a criação de modelos devidamente enquadrados nos segmentos alvo. O processo de identificação seria baseado num conjunto de perguntas (variando desde das características sócio-demográficas a perguntas que procuravam avaliar o comportamento perante o risco e os objetivos de investimentos). O BBP tinha a noção que a importância das perguntas variava com as respostas que iam sendo dadas. A relevância de uma resposta dada a uma determinada pergunta variava com o contexto definido nas perguntas anteriores. (SANTOS, 2002)

2.3.4- Técnicas de Data Mining.

a) Inteligência Artificial.

Tentando imitar a maneira de pensar do ser humano a Inteligência Artificial, como embasamento estatístico, tentar recuperar informações que podem auxiliar na descoberta de novas interações entre os dados. Para tal, esta técnica heurística, exige mais processamento, o que só foi conseguido, com um bom custo/benefício, na década passada, resultando em seu maior aproveitamento nos SGBDs, e em seus módulos de otimização de consulta (LOYOLLA, 2002).

b) Redes Neurais Artificiais.

Uma Rede Neural Artificial (RNA) é uma técnica computacional que constrói um modelo matemático, emulado por computador, de um sistema neural biológico simplificado, com capacidade de aprendizado, generalização, associação e abstração. A rede neural artificial tenta aprender padrões diretamente dos dados através de um processo de repetidas apresentações dos dados à rede, ou seja por experiência. Dessa forma, uma rede neural artificial procura por relacionamentos, constrói modelos automaticamente, e os corrige de modo a diminuir seu próprio erro.(AURELIO, 2002).

c) Algoritmos Genéticos

São métodos computacionais de busca baseados nos mecanismos de evolução natural e na genética. Em Algoritmos Genéticos, uma população de possíveis soluções para o problema em questão evolui de acordo com operadores probabilísticos concebidos a partir de metáforas biológicas, de modo que há uma tendência de que, na média, os indivíduos representem soluções cada vez melhores à medida que o processo evolutivo continua (TANOMARU, 1995).

d) Lógica indutiva.

A lógica indutiva (LI) pode ser vista como uma máquina de aprendizado em lógica de primeira ordem, onde as relações são apresentadas no contexto de banco de dados dedutivos. A Lógica Indutiva é relevante para a descoberta de conhecimento em base de dados relacionais e dedutivas, pois pode descrever padrões envolvendo mais de uma relação (AURELIO, 2002).

e) Métodos Estatísticos

Os métodos estatísticos são responsáveis pela descrição dos dados armazenados das bases de dados, através de técnicas como: análise de discriminantes, ajustes de curvas, análise de variância, testes de hipóteses, cálculos de probabilidades e outros mais. Apesar de todas essas técnicas você só pode retirar melhor proveito dessas ferramentas com a presença de um profissional para o auxílio na retirada das informações (AURELIO, 2002).

f) Análise de Cluster

Esta técnica agrupa informações homogêneas de grupos heterogêneos entre os demais e aponta o item que melhor representa cada grupo, permitindo, desta forma, que consigamos perceber a característica de cada grupo. (DAL'ALBA, 2002).

g) Regressão

As técnicas de regressão auxiliam a prever o futuro, baseado no passado. Para isso utiliza-se de dados históricos. Existem várias técnicas de regressão e deve-se utilizar a mais adequada em cada caso (DAL'ALBA, 2002).

h) Árvore de decisão.

Nesta técnica você escolhe a variável que quer avaliar e o software procura as mais correlacionadas e monta a árvore com várias ramificações. As árvores de decisão são meios de representar resultados de Data Mining na forma de árvore, e que lembram um gráfico organizacional horizontal. Dados um grupo de dados com numerosas colunas e linhas, uma ferramenta de árvore de decisão pede ao usuário para escolher uma das colunas como objeto de saída, e aí mostra o único e mais importante fator correlacionado com aquele objeto de saída como o primeiro ramo (nó) da árvore de decisão. Isso significa que o usuário pode rapidamente ver qual o fator que mais direciona o seu objeto de saída, e o usuário pode entender porque o fator foi escolhido. Uma boa ferramenta de árvore de decisão vai, também, permitir que o usuário explore a árvore de acordo com a sua vontade, do mesmo modo que ele poderá encontrar grupos alvos que lhe interessem mais, e aí ampliar o dado exato associado ao seu grupo alvo. Os usuários podem, também, selecionar os dados fundamentais em qualquer nó da árvore, movendo-o para dentro de uma planilha ou outra ferramenta para análise posterior. Nesta técnica consegue-se saber os itens que mais influenciam uma determinada variável. Por exemplo, consegue-se saber que tipo de cliente está insatisfeito. (GIMENES, 2002).

2.3.5 - O Processo de Implementação do Data Mining.

Todo o Processo de implementação do Data Mining se dá através da análise do problema de uma determinada organização com o objetivo geral de solucionar problemas tradicionalmente difícil. Para a execução de projetos de Data Mining e a correta alocação de recursos para os mesmos. A SPSS, empresa especializada em prover soluções de Business

Intelligence com ênfase em data mining, acaba de disponibilizar na internet uma metodologia para o processo de mineração de dados. Trata-se do CRISP-DM (Cross Industry Standard Process for Data Mining), um projeto que padroniza conceitos e técnicas na busca de informações em bancos de dados. A metodologia dividiu o processo de implementação de um sistema de Data Mining em seis fases interdependentes para que o mesmo atinja seus objetivos finais. São elas:

Entendimento do negócio:

O entendimento do negócio se dar na análise de todo o problema da organização, avaliando suas metas, perspectivas e objetivos a serem alcançados com o projeto Data Mining, convertendo tudo isso para a aplicação do Data Mining criando com isso um plano inicial a resolução dos problemas encontrados.

Entendimento dos dados:

No entendimento de dados se encontra a fase da retirada de uma amostra de dados da base de dados para serem avaliados e na avaliação do ambiente em que esses dados se encontram armazenados.

Preparação dos dados:

Na preparação dos dados se encontra a fase de criação de programas de extração de dados, a limpeza dos dados retirando todas as informações desnecessárias e transformação dos dados para o uso das técnicas Data Mining.

Modelagem:

Na modelagem se dar o processo de seleção do(s) algoritmo(s) ou das técnicas a serem utilizadas no processo de retirada dos dados das bases de dados

Avaliação do modelo:

Ao final da fase de modelagem, vários modelos devem ter sido avaliados sob a perspectiva do analista responsável. Agora o objetivo passa a ser avaliar os modelos com a visão do negócio, se certificando que não existem falhas ou contradições com relação às regras do negócio.

Publicação:

A criação e validação do modelo permitem avançarmos mais um passo, no sentido de tornar a informação gerada acessível. Isto pode ser feito de várias maneiras, desde a criação de um software específico para tal, até a publicação de um relatório para uso interno.

O Objetivo geral de todo o processo de implementação é facilitar todo o processo de implementação do sistema a ser utilizado pela organização, gerando com isso confiabilidade em todo o processo de retirada de informação e assegurando que todo o processo está sem falhas.

2.5.6 - Algumas Barreiras ao uso do Data Mining.

Como já visto ao longo do trabalho a implementação do Data Mining em bases de dados passar por inúmeros processos, desde da escolha da técnica a ser utilizada há ferramenta que melhor se adequar ao processo de mineração de dados para o auxílio aos Sistemas de Apoio à Decisão (SAD). Como se sabe para toda nova tecnologia existe barreiras a serem enfrentadas para a popularização das mesmas e segundo SERRA (2002) as barreiras são as seguintes:

Altos Custos.

O Alto custo da maioria das ferramentas faz com que fique difícil a disseminação das mesmas entre as organizações. Uma simples operação matemática mostra que um projeto, com o custo de US\$20.000 por usuário, pode atender não mais que um seleto grupo de usuário. Mais com a popularização dessas ferramentas já vem surgindo fornecedores com produtos com custo mais baixo, mais ainda é necessário a redução mais ainda desses custos para que possa atingir todos os níveis de usuários.

Necessidades de grandes volumes de dados

Um dos maiores barreiras para a utilização do Data Mining no passado era a necessidade de se armazenar grandes volumes de dados para que com isso fosse possível a mineração de dados com as ferramentas existentes. Hoje em dia as ferramentas dar o usuário a possibilidade de minerar porções de dados em seu próprio computador, o que pode, efetivamente, contornar essa barreira. Embora não tenham a intenção de substituir aplicações que necessitam de grandes volumes de dados, essas ferramentas podem prover um acesso alternativo que pode também ser usado em conjunto com as ferramentas 'pesadas'.

3. Complexidade das ferramentas

Uma outra barreira e bem mais complexa é a complexidade das ferramentas existentes no mercado para o processo de mineração de dados. Pois muitas ferramentas fazem os seus trabalho em uma 'caixa-preta', não permitindo que se saiba como seus resultados foram alcançados. Apesar de existir técnicas de Data Mining mais compreensíveis, como, por exemplo, as árvores de decisão, que permitem que qualquer um, com conhecimento básico de computadores, possa utilizar e compreender o processo.

Contudo, o poder dessas ferramentas não chega nem perto daquelas que utilizam técnicas sofisticadas.

Dificuldade na preparação dos dados para a mineração

Existe uma grande dificuldade na preparação dos dados para a mineração de dados pois existe muitas e trabalhosas tarefas em um projeto de Data Mining, sendo considerado 80 por cento do trabalho. Mas tarefas a serem realizadas para o processo de Mineração de Dados existem a tarefa de limpeza de dados para que não haja erros lógicos, a centralização dos dados em um único local e por fim a escolha dos dados certos para o processo de mineração.

Dificuldade de realizar a análise custo/benefício do projeto de Data Mining

Não se sabe ao certo se o projeto de Data Mining vai ser útil para a tomada de decisão das organizações até que ele seja concluído e que se faça a análise das informações obtidas. Pois o projeto de Data Mining faz com que o usuário do Data Mining reconheça as tendências ou padrões das organizações. Valores fora do comum em um grupo de dados fazem com que seja mais difícil minerar.

Viabilidade dos fornecedores de ferramentas Data Mining.

Como se sabe existe um crescimento gigantesco das tecnologias em todo o mundo e que a cada dia surge novas tendências e tecnologias e que outras vão ficando fora de uso. Uma das principais barreiras encontrada pelas organizações é a viabilidade dos fornecedores de ferramentas Data Mining, pois as organizações procuram ferramentas confiáveis para hoje e para o futuro. Sendo um dos passos mais importante para o projeto de Data Mining a escolha do fornecedor e da ferramenta adequada.

3. ANÁLISE DOS RESULTADOS

A pesquisa foi realizada em quatro etapas, que são: pesquisa bibliográfica, levantamento de dados, análise das informações e intervenção.

O levantamento bibliográfico teve como principal tarefa a pesquisa de materiais científicos e comerciais para a elaboração do trabalho, a mesma foi realizada em livros e nos sites da internet com o intuito de levantar material atualizado para a conclusão do trabalho.

Para o tratamento de dados foi realizado um levantamento dos softwares existentes no mercado para a realização da mineração. Foram Escolhidos os softwares Wizrule e SuperQuery, pelo fato de se tratar de um software de interação amigável, baixo custo e que funciona em plataformas padrão. Além de disponibilizarem cópias de demonstração para análise. Descreveremos a utilização em detalhes dos softwares acima citados:

Os dados escolhidos encontram-se no banco de dados da Biblioteca Henrique Castriciano localizada na FARN. Foram selecionadas algumas tabelas que foram julgadas necessárias, fazendo a carga delas em planilhas do Excel2000 e tabelas do Access2000.

No Banco de dados estudado continha tabelas com informações armazenadas que serviu de base para a Mineração desses dados, nestas tabelas continham informações do tipo:

- Cadastrados de usuários
- Acervos de livros, dissertações, teses e periódicos.
- Informações sobre os empréstimos realizados.

Para a realização do trabalho foi necessário a realização de uma pesquisa qualitativa do tipo pesquisa-ação. Qualitativa pelo fator do ambiente natural ser a fonte direta para a coleta de dados e o pesquisador ser o instrumento-chave e é de natureza descritiva porque envolve o uso de técnicas padronizadas de coleta de dados e observação sistemática. Assumindo a forma de levantamento. Do tipo pesquisa-ação pelo fato de conter traços essenciais, como a análise, coleta de dados e conceituação dos problemas; planejamento da ação, execução e nova coleta de dados para avaliá-la.

Vejamos agora os requisitos mínimos para a utilização do Wizrule e o SuperQuery:

Tabela3- Requisitos mínimos para a utilização das ferramentas Wizrule e SuperQuery.

Wizrule	SuperQuery
❖ 486DX Microprocessador	❖ 486DX Microprocessador
❖ 8MB Ram (18MB recomendado)	❖ 8MB Ram (18MB recomendado)
❖ Windows 95/98/Me/2000/XP	❖ Windows 95/98/Me/2000/XP

3.1- WIZRULE

O Wizrule é um software fabricado pela companhia Wizsoft, para que ele nos auxilia na obtenção das tendências de comportamentos dos usuários da biblioteca.

Um fator essencial para a escolha desta ferramenta foi a possibilidade de se fazer um download de um demonstrativo da ferramenta Wizrule na versão 4.01 que trabalha com no máximo 1000 registros fazendo associações entre eles.

Um outro fator e de muita importância para a escolha desta ferramenta para a mineração foi o fato de ser um software muito amigável. Pois uma pessoa que só tenha o conhecimento básico de regras de inferência, ou simplesmente lógica matemática, pode utilizar esta ferramenta sem maiores transtornos. Ele permite a realização de uma análise básica nos dados sem exigir grande interferência do usuário.

A desvantagem encontrada no uso dessa ferramenta é fato de que ela tem a limitação de analisar somente uma base de dados de cada vez, não sendo possível a visualização das regras em diferentes períodos.

Os resultados encontrados com a utilização da ferramenta Wizrule foram pesquisados utilizando o universo de 1000 registros do banco de dados. Os dados apresentados contem probabilidade superior a 60 %, que é o valor mínimo aceito pelo software. Vejamos agora os resultados:

- No Curso de Direito cerca de 61,9% dos usuários que fazem empréstimos na biblioteca é do sexo masculino.
- No mês de setembro cerca de 96,2 % das devoluções realizadas foram entregues no prazo.
- Em setembro no curso de Direito, cerca de 61,9% dos empréstimos foram feitas por alunos do sexo masculino.

- Em setembro no curso Ciências Contábeis, cerca de 60% dos empréstimos foram feitas por do sexo masculino.
- Em novembro no curso de Direito, cerca de 67,3% dos empréstimos foram feitas por do sexo masculino.
- Em outubro dos alunos de Licenciatura em Computação que retiraram livros, cerca de 62,9%, foram do sexo masculino.

Conclui-se que no curso de Direito os alunos são os usuários mais freqüentadores da biblioteca da FARN e os alunos do sexo masculino em particular. Também certifica-se que ao fim de cada ano letivo intensifica-se a procura de material na biblioteca com já era de se esperar.

3.2- SUPERQUERY

O SuperQuery é um software desenvolvido pela AZMY Thinkware que é uma firma especializada no desenvolvimento de ferramentas de análise de dados. O superQuery é uma ferramenta de análise de descoberta de dados que roda em plataforma Windows, usado tecnologia de regras de indução. O SuperQuery lê dados em Access, xBASE, Excell, Paradox e diretamente de arquivos textos, podendo também usar o ODBC para a maioria dos banco de dados populares que tem drivers de ODBC.

Os resultados encontrados com a utilização da ferramenta SuperQuery, foram pesquisados utilizando o universo de 6682 na tabela empréstimo e 994 na tabela usuário registros do banco de dados. Vejamos agora os resultados:

Tabela 1- Valores estatísticos para material contido na biblioteca da FARN.

Ano de Publicação	Quantidade	%
1998	471	5
2001	376	4
2000	349	3
1999	337	3
1997	280	3
1978	250	3
1976	212	2

A tabela acima apresenta os sete maiores valores com relação ao ano de publicação do acervo da biblioteca. Nota-se que a maior concentração de publicações está relacionada aos exemplares publicados nos últimos quatro anos.

Tabela2- Valores estatísticos das edições contidas no acervo da biblioteca da FARN.

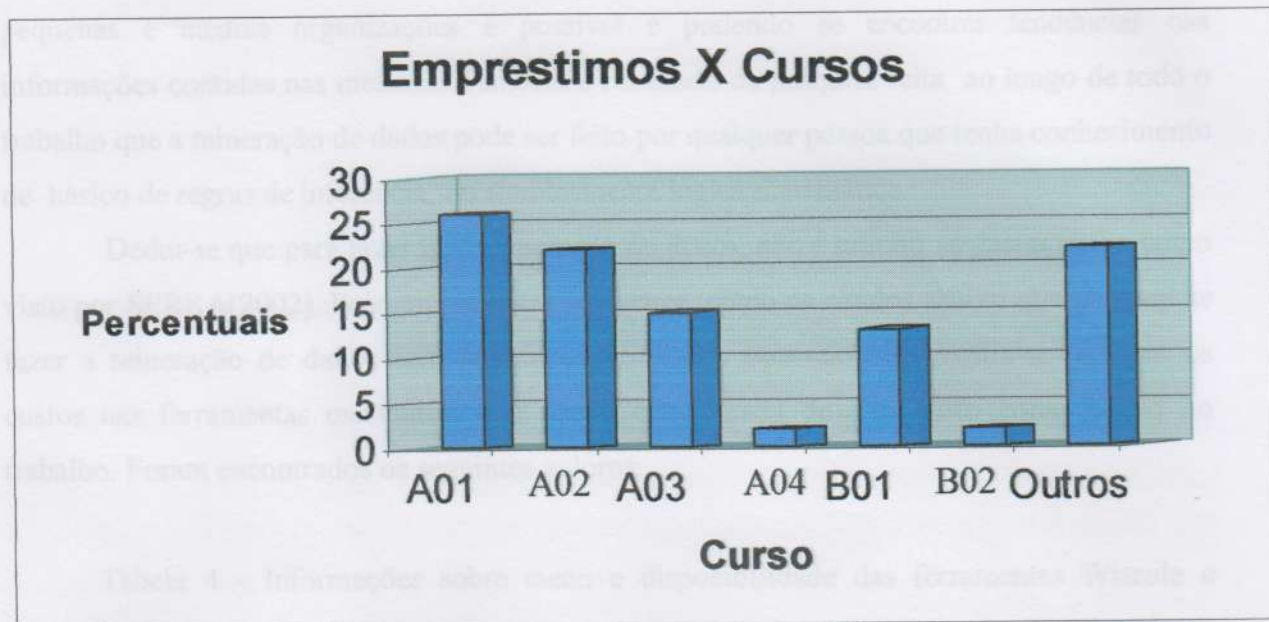
Editores	Quantidade	%
Forense	604	6
Saraiva	590	6
Atlas	511	5
Freitas Bastos	340	3
Revistas dos Tribunais	299	3
LTR	170	2

A tabela acima apresenta os sete maiores valores com relação aos Editores de exemplares contidos no acervo da biblioteca. Verifica-se que a maior concentração de publicações encontradas foram as citadas acima. Os resultados dessas informações servem para uma futura negociação.

- Dos alunos cadastrados na biblioteca, somente 8% dos usuários que realizam empréstimos é do curso bacharelado em Sistema de Informação e são do sexo feminino.
- Dos alunos cadastrados na biblioteca, somente 1% dos usuários são do sexo feminino é do curso de licenciatura.
- Dos alunos cadastrados na biblioteca, somente 2% dos usuários é do curso do sexo masculino é do curso de licenciatura.

O gráfico abaixo mostra o resultado encontrado com a ferramenta SuperQuery, utilizando a tabela empréstimos para a mineração de dados.

4- Gafico1- Representação gráfica dos resultados encontrados com a mineração de dados referentes as tabelas: Empréstimo X Curso.



Legenda:

A01- Administração.

A02- Ciências Contábeis.

A03- Direito.

A04-Marketing.

B01- Bacharelado em Sistema de Informação.

B02- Licenciatura em Computação.

4 - CONCLUSÃO

Conclui-se por meio desse trabalho que a utilização da mineração de dados em pequenas e médias organizações é possível e podendo se encontrar tendências nas informações contidas nas mesmas. Também é resultado da pesquisa feita ao longo de todo o trabalho que a mineração de dados pode ser feito por qualquer pessoa que tenha conhecimento de básico de regras de inferência, ou simplesmente lógica matemática.

Dedui-se que para fazer uma mineração de dados, não é preciso se gastar muito como visto por SERRA(2002). Pois encontramos softwares, como os citados abaixo que possível se fazer a mineração de dados sem grandes dificuldades operacionais. Avalia-se também os custos nas ferramentas existentes, que não é considerada de alto custo como citado no trabalho. Foram encontrados os seguintes valores:

Tabela 4 - Informações sobre custo e disponibilidade das ferramentas Wizrule e SuperQuery.

Wizrule	SuperQuery
US\$ 1395,00	US\$ 450,00
Disponível: www.wizrule.com	Disponível: www.azmy.com/

As tendências encontradas com a mineração dos dados foram as seguintes:

- Uma grande quantidade existente de alunos do sexo feminino cadastradas na biblioteca o que corresponde a 55% do total dos usuários, 37% alunos do sexo masculino e 8 % que não registraram sexo.
- Uma outra tendência encontrada é que apesar da grande quantidade de mulheres existentes nos cadastros da biblioteca, identificou que a quantidade de homens que utilizam-se dos serviços das mesmas são bem maiores que o números de mulheres e em especial os usuários do curso de Direito. Mostrando uma tendência dos homens a serem mais leitores que as mulheres.
- Encontrou-se também que os livros disponíveis no acervo da biblioteca foram publicados nos últimos quatro anos, mostrando a tendência da biblioteca a estar sempre com publicações atualizadas. Um outro ponto interessante é que foi encontrada uma tendência pelas editoras: Forense, Saraiva, Atlas, Freitas

Bastos e Revistas dos Tribunais. Que com este resultado poderia servir para a negociação com as editoras para uma futura transação na aquisição de novos produtos. Sendo isso mais um exemplo de aplicações Data Mining para o comércio, provando que se é possível a mineração de dados e a obtenção de resultados em pequenas e médias organizações.

- Conclui-se também com este trabalho que é possível a elaboração de pesquisa em qualquer área de interesse, com obtenção de resultados favoráveis para uma possível a decisão de negócios, ou até mesmo na previsão futura de tendências ou comportamentos de futuros clientes.

4.1- SUGESTÃO

Sugeri-se para trabalhos futuros a continuação dessa pesquisa o estudo das técnicas existentes, tais como: redes neurais, algoritmos genéticos, métodos estatísticos, regressão, análise de cluster, lógica indutiva e inteligência artificial, qual a melhor se adequa a uma área específica do negócio.

[7] DATA Mart. Disponível <<http://www.datawarehouse.info.br>>. Acesso em 10 de setembro de 2002.

[8] DALFOPO, Gualdo & GRIPA, Nelson. Data Warehouse. Disponível <<http://www.datawarehouse.com>>. Acesso em 04 de abril de 2002.

[9] GIMENES, Eduardo. Introdução. Monografia Apresentada para Graduação em Processamento de Dados - Peter Tapalunga. Disponível <<http://apccia.rebo.com.br/diagramas>>. Acesso em 04 de agosto de 2002.

[10] GOLDBERG, D. P. Genetic algorithms in search, optimization, and machine learning. Addison-Wesley Publishing Company Inc., 1989.

[11] INTRODUÇÃO à mineração de dados, descobrindo valor escondido nos seus dados de dados. Disponível <<http://www.marketingdiretoria.com.br>>. Acesso em 01 de abril de 2002.

[12] JUNIOR, Omar de Oliveira Braz. Agrado para Auxiliar na construção de convergência e definição de intervalos ideais para parâmetros em algoritmos genéticos. Disponível <<http://www.inf.ufes.br/~josecarlos/otima/Omar/junior.htm>>. Acesso em 01 de agosto de 2002.

[13] AURELIO, Marco VELLASCO, Marley, LOPES, Carlos Henrique. Base de conhecimento e mineração de dados. ICA - Laboratório de inteligência computacional aplicada do Departamento de Engenharia Elétrica, FURG - Rio, 2002.

5- REFERÊNCIAS BIBLIOGRÁFICAS

- [1] ANDRINI, Yeda Maria Heineck; ZOMER, Silvania da Silva. **Inteligência Empresarial Estratégica na Sociedade do Conhecimento**. Disponível < www.informacaoesociedade.ufpb.br/919901.pdf > Acesso em 10 de novembro de 2002.
- [2] ASSUNÇÃO, Luis. **Data Warehousing**. Disponível < <http://www.assuncao.eti.br/luisdwh.htm> > Acesso em 18 de novembro de 2002.
- [3] BR-BUSINESS. **BI - Business Intelligence**. Disponível < <http://www.br-business.com.br/brb/busintel.htm> > Acesso em 22 de novembro de 2002.
- [4] BRUSSO, Marcos José. **O uso de mineração de dados na descoberta do comportamento do usuário da web**. Disponível < <http://www.inf.ufrgs.br/pos/SemanaAcademica/Semana99/brusso/brusso.html> >. Acesso em 02 de junho de 2002.
- [5] CID, Dante Jose Alexandre; PASSOS, Emmanuel P. L. **Análise de ferramentas Data Mining**. Disponível < www.ica.ele.puc-rio.br/cursos/download/DM-xprule.pdf > Acesso em 10 de novembro de 2002..
- [6] DAL'ALBA, Adriano. **Introdução**. Disponível www.geocities.com/SiliconValley/PORT/5072/. Acesso em 10 de setembro de 2002.
- [7] **DATA Mart**. Disponível <<http://www.datawarehouse.inf.br>>. Acesso em 10 de setembro de 2002.
- [8] DALFOFO, Osvaldo & GRIPA, Robson. **Data Warehouse**. Disponível <<http://www.datawarehouse.com>> Acesso em 04 de abril de 2002.
- [9] GIMENES, Eduardo. **Introdução**, Monografia Apresenta para Graduação em Processamento de Dados - Fatec Taquaritinga, Disponível <<http://geocities.yahoo.com.br/dugimenes>>. Acesso em 04 de agosto de 2002.
- [10] GOLDBERG, D. E. **Genetic algorithms in search, optimization, and machine learning**. Addison-Wesley Publishing Company Inc., 1989.
- [11] **INTRODUÇÃO à mineração de dados, descobrindo valor escondido nas suas bases de dados**. Disponível <<http://www.marketingdeprecisao.com.br>> Acesso em 01 de abril de 2002.
- [12] JUNIOR, Osmar de Oliveira Braz. **Agentes para Auxiliar nas condições de convergência e definição de intervalos ideais para parâmetros em algoritmos genéticos**. Disponível < <http://www.inf.ufsc.br/~jbosco/adriana/Osmar/artiagen.htm> > Acesso em 27 de agosto de 2002.
- [13] AURELIO, Marco; VELLASCO, Marley; LOPES, Carlos Henrique. **Descoberta de conhecimento e mineração de dados**. ICA - Laboratório de inteligência computacional aplicada do Departamento de Engenharia Elétrica, PUC - Rio, 2002.

- [14] LOYOLLA, Waldomiro - **Avaliação do estado da arte e produtos Data Mining**. Dissertação de mestrado de informática da Universidade Católica de Brasília.
- [15] NAVEGA, Sérgio. **Princípios essenciais do Data Mining**. Disponível <www.intelliwise.com/snavega> Acesso em 28 de outubro de 2002.
- [16] SANTOS, José; HENRIQUES, Nuno A. C.; REIS, Vanda. **Data mining/Data Warehousing**. Disponível <<http://students.fct.unl.pt/users/nach/dmdw/>>. Acesso em 08 de novembro de 2002.
- [17] SERRA, Laércio. **A essência do Business Intelligence**. São Paulo. Editora Berkeley. 2002
- [18] TANOMARU, J. **Motivação, fundamentos e aplicações de algoritmos genéticos**. Congresso Brasileiro de Redes Neurais, 1995, Curitiba.
- [19] **TÉCNICAS, de Data Mining**. Disponível <<http://geocities.yahoo.com.br/dugimenes/tecnicas.htm>> Acesso em 20 de novembro de 2002.
- [20] TORRES, João Vitor. **Bases de dados na educação**. 2ª Versão 2002.
- [21] TREICHEL, Davi. **Ferramenta para visualização de resultados gerados por Data Mining**. Disponível <atlas.ucpel.tche.br/~projgrad/2002-1/antedavi.htm>. Acesso em 08 de novembro de 2002.
- [22] VASCONCELOS, Marco A Sanford de; VASCONCELOS, Aymara Heineck de. **Inteligência estratégica empresarial um passaporte para o futuro**. Dissertação do mestrado do Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina, 2002.
- [23] ZANASI, Alessandro - **Inteligência competitiva usando o Data Mining em bases públicas**, Disponível <<http://open.cineca.it/datamining>>, acesso em 04 de abril de 2002.
- [24] **QUÊ É O QUE É Data Mining?** Disponível <<http://www.parcos.com.br/contatos/datamining.htm>> Acesso em 04 de março de 2002.
- [25] O'BRIEN, James A. **Sistemas de Informação e as Gerências gerenciais na era da Internet**. Editora Saraiva, 2001.
- [26] PASSOS, Daniel José Alexandre Cid e Emmanuel P. L. **Análise de Ferramentas de Data Mining**. Disponível <www.lis.uec.br/curriculo/download/DM-wiz> Acesso em 14 de outubro de 2002.
- [27] **QUEM procura não encontra** de informação podem encontrar muito mais do que se busca de minerar os dados. <<http://www.kdnugest.com/presentations/2002.html>> Acesso em 04 de março de 2002.

6. BIBLIOGRAFIA CONSULTADA

- [24] ARAÚJO, Chiarlan Emerciriano Silva. **Banco de Dados em sistema de vídeo sob demanda**. Dissertação do relatório de graduação da Universidade Federal do Rio Grande do Norte.
- [25] **Data Mining**. Disponível <http://www.datawarehouse.inf.br/datamining.asp>. Acesso em 04 de março de 2002.
- [26] **Data Mining**. Trabalho apresentado na disciplina de Banco de Dados, da Monografia em Formação de Analistas de Sistema de Informação da UERJ, Disponível em <http://hps.infolink.com.br/mpolito/mining/mining.htm>. Acesso em 04 de maio de 2002.
- [27] **Data Mining/ Data Warehouse**. Disponível <<http://students.fct.unl.pt.users/jms/dmdv/datamining/>> Acesso em 04 de março de 2002.
- [28] FREITAS, Alex A. **Mineração busca informações estratégicas em grandes bancos de dados**. Disponível <<http://www.cesar.org.br/analise/n-32/fra-mineração.html>> Acesso em 20 de maio de 2002.
- [29] **If - organizing Data Mining e seus algoritmos**. Disponível <<http://black.rc.unesp.br/ccompli.ia/ia2000/datamining/dm-self-organizing%20data%20mining%20e%20seus%20algoritmos.htm>> Acesso em 04 de abril de 2002.
- [30] **INFORMÁTICA Brasileira em Análise**. Disponível <http://www.cesar.org.br/analise/n_32/n_32.html> Acesso em 04 de março de 2002.
- [31] **INTELIGÊNCIA Obtida pela aplicação de Data Mining em bases de teses Francesas sobre o Brasil**. Disponível <<http://www.ibict.br/cionline/300201/3020104.pdf>> Acesso em 01 de março de 2002.
- [32] MARCOS, José Brusso. **O uso de mineração de dados na descoberta do comportamento do usuário da Web**. Disponível <<http://www.lch.dcc.ufmg.br/publicacoes/mineracao.html>> Acesso em 06 de agosto de 2002.
- [33] **O QUE é Data Mining?** Disponível <<http://www.gemco.com.br/conceitos/datamining.htm>> Acesso em 04 de março de 2002.
- [34] O'BRIEN, James A. **Sistemas de informação e as decisões gerenciais na era da Internet**. Editora Saraiva, 2001.
- [35] PASSOS, Dante José Alexandre Cid e Emmanuel P. L.; **Análise de Ferramentas de Data Mining**. Disponível www.ica.ele.puc-rio.br/cursos/download/DM-wiz Acesso em 18 de outubro de 2002.
- [36] **QUEM procura Acha montanhas de informação podem esconder minas de ouro. É hora de minerar os dados**. <<http://www.kdnugget.com/press/exame2002.html>> Acesso em 04 de março de 2002.

[37] STAIR, Ralph M. **Princípios de sistemas de informação. Uma abordagem gerencial.** 2ªed. Rio de Janeiro: LTC Editora. 1998.

[38] **WMS: A solução que preserva o lucro do gestor** Disponível
<<http://www.gemco.com/conceitos/wms.htm>> Acesso em 04 de março de 2002.