DATA MINING IN SUPPORT OF THE KNOWLEDGE MANAGEMENT

Eric Rommel Galvão Dantas (Universidade Federal de Pernambuco, PE, Brasil) – ergd@cin.ufpe.br
Ryan Ribeiro de Azevedo (Universidade Federal de Pernambuco, PE, Brasil) – rra2@cin.ufpe.br
José Carlos Almeida Patrício Júnior (Centro Universitário de João Pessoa, PB, Brasil) - jcapjunior@hotmail.com
Daniel Silva de Lima (Centro Universitário de João Pessoa, PB, Brasil) - danielpk.lima@gmail.com

Wendell Campos Veras (Universidade Federal de Pernambuco, PE, Brasil) – wcv@cin.ufpe.br

Abstract: The process of knowledge management demand a constant number of decisions about the activities in an organization, based on these decisions to the knowledge gained with experience, directly impacting on the entire productive chain of the organization. To assist in that management, computational tools for generating new knowledge are essential. This article describes the process of knowledge discovery in databases, through the mining of data, demonstrating its applicability in the process of support in decision making.

Keywords: Knowledge Management, Knowledge Discovery, KDD, Data Mining

MINERAÇÃO DE DADOS NO APOIO A GESTÃO DO CONHECIMENTO

Resumo: O processo de gestão do conhecimento demanda um constante número de decisões acerca das atividades presentes em uma organização, tendo como base para essas decisões o conhecimento adquirido com a experiência profissional, impactando diretamente em toda a cadeia produtiva da organização. Para auxiliar nesse gerenciamento, ferramentas computacionais geradoras de novos conhecimentos são essenciais. Este artigo descreve o processo de descoberta de conhecimento em bases de dados, através da mineração de dados, demonstrando suas aplicabilidades no processo de apoio na tomada de decisão.

Palavras-chave: Gestão do Conhecimento; Descoberta de Conhecimento; KDD; Mineração de dados.

1. INTRODUÇÃO

A informação é um dos ativos mais importante para os negócios das organizações, tornando-se algo essencial para ganho de competitividade entre as empresas de pequeno, médio e grande porte. As estratégias assumidas para tal ganho devem basear-se em informações concretas, visando uma minimização na ocorrência de erros para a tomada de decisões por parte dos gestores.

Avanços tecnológicos têm facilitado a obtenção dessas informações através de processos de *Knowledge Discovery in Database (KDD*), ou seja, Descoberta de Conhecimento em Banco de Dados. O KDD pode ser visto como o processo de descoberta de padrões e tendências por análise de grandes conjuntos de dados, tendo como principal etapa o processo de mineração, consistindo na execução prática de análise e de algoritmos específicos que, sob limitações de eficiência computacionais aceitáveis, produz uma relação particular de padrões a partir de dados FAYYAD *et al.* (1996).

Tal processo identifica padrões e descobre informações relevantes que auxiliam na formação de posturas estratégicas de *marketing*, na busca e conquista de clientes, na descoberta de falhas em linhas de produção, entre outras. Nesse processo, na etapa de mineração de dados, algoritmos têm a função de gerar regras de associação, descrevendo assim, padrões de relacionamento entre itens de uma base de dados.

Este trabalho realiza uma análise das regras geradas, assim como as evoluções propostas e os aspectos da integração destas tecnologias com o KDD, conceituando, esclarecendo e despertando o interesse das empresas sobre a importância da sua utilização na tomada de decisões estratégicas para o negócio.

Este artigo está organizado da seguinte forma: na Seção 2 são apresentados os conceitos a cerca do entendimento da gestão do conhecimento; na Seção 3 são vistos os conceitos e benefícios do processo de descoberta de conhecimento KDD, mineração de dados e regras de associação. São apresentados na Seção 4 a metodologia e os resultados. Na Seção 5 são apresentados os trabalhos relacionados. Por fim, na Seção 6 apresentam-se as conclusões e trabalhos futuros.

2. GESTÃO DO CONHECIMENTO

O valor de uma decisão estratégica para os negócios depende das informações disponíveis ao gestor de uma organização, da capacitação que este possui de interpretá-las e da experiência para associá-las de maneira conveniente. DAVENPORT e PRUSAK (1998) acrescenta que a única vantagem sustentável de uma empresa é o que ela coletivamente sabe, a eficiência com que ela usa o que sabe e a prontidão com que ela adquire e usa novos conhecimentos.

Informação é o resultado do processamento de dados num formato que tem significado para o usuário respectivo e que tem valor real ou potencial nas decisões presentes ou prospectivas DAVIS (1974). O conhecimento necessário para se decidir e/ou avaliar torna-se disponível por meio de informações SANCHES (1997). A gerência da informação ou a administração do conhecimento e da informação é uma breve definição de Gestão do Conhecimento (GC). Segundo DAVENPORT e PRUSAK (1998), a GC pode ser vista como uma série de ações gerenciais constantes e sistemáticas que facilitam os processos de

criação, registro e compartilhamento do conhecimento nas organizações.

Para KRUGLIANSKAS e TERRA (2003), GC significa organizar as principais políticas, processos e ferramentas gerenciais e tecnológicas à luz de uma melhor compreensão dos processos de geração, identificação, validação, disseminação, compartilhamento, uso e proteção dos conhecimentos estratégicos para gerar resultados (econômicos) para a empresa e benefícios para os colaboradores internos e externos (*stakeholders*).

ALVARENGA NETO (2005) explica ainda que GC é o conjunto de atividades voltadas para a promoção do conhecimento organizacional, possibilitando que as organizações e seus colaboradores possam sempre se utilizar das melhores informações e dos melhores conhecimentos disponíveis, com vistas ao alcance dos objetivos organizacionais e maximização da competitividade.

Para apoiar a GC, são utilizadas diversas técnicas da Inteligência Computacional, tais como, Redes Neurais Artificiais, Lógica *Fuzzy*, e métodos estatísticos. Estas técnicas demandam altos investimentos em *software* e *hardware*, não menos ainda em pessoal capacitado. WU e WANG (2006) comentam que grande parte dos investimentos em GC é destinada aos Sistemas de Gestão do Conhecimento (SGC), ferramentas baseadas na Tecnologia da Informação (TI) capazes de suportar os processos de criação, armazenamento, recuperação, transferência e aplicação do conhecimento. A Seção 3 abordará o processo de descoberta de conhecimento.

Evidentemente apenas os recursos tecnológicos, não garantirão uma perfeita GC. Faz-se necessário um capital humano bem treinado e com experiência de mercado para interpretar as informações e conhecimentos disponibilizados pela tecnologia para assim, tomar as melhores decisões. Estas atitudes são essenciais mesmo quando suas práticas limitam-se a não decidir, visto que a falta de decisão, por si só, já é uma forma de decisão.

3. DESCOBERTA DE CONHECIMENTO EM BASES DE DADOS (KDD)

A extração do conhecimento é uma área dinâmica e evolutiva, envolvendo integrações com outras áreas de conhecimento como Estatística, Inteligência Artificial e Banco de Dados. Os padrões extraídos devem ser além de confiáveis, compreensíveis e úteis, podendo empregar o conhecimento com utilidade e tirar proveito de alguma vantagem, seja científica ou comercial. De acordo com FAYYAD *et al.* (1996), o processo de KDD é constituído de diversas fases, explicadas na Seção 3.1, e tem início na análise do entendimento do domínio da aplicação e dos objetivos a serem realizados.

3.1. Fases do KDD

Após a fase inicial, o foco passa a ser a escolha ou seleção da massa de dados a ser minerada, podendo ser um conjunto de dados ou um subconjunto de variáveis onde a extração será realizada. A fase de *Data Cleaning* e Pré-Processamento tem por objetivo assegurar a qualidade dos dados envolvidos no KDD realizando operações básicas como a remoção de ruídos, que podem ser, por exemplo, atributos nulos. A fase seguinte consiste na Seleção e Transformação dos dados em que serão selecionados os atributos realmente interessantes ao usuário, além de transformados utilizando o padrão ideal para aplicar algoritmos de mineração.

Após a realização das fases anteriores, a Mneração de Dados (*Data Mining*) é iniciada. Esta fase é a mais importante do KDD, sendo realizada através da escolha do método e do algoritmo mais compatível com o objetivo da extração, a fim de encontrar padrões nos dados que sirva de subsídios para descobrir conhecimentos ocultos. Na Seção 3.2 um aprofundamento desta fase será realizado.

A Avaliação ou Pós-Processamento é a fase que identifica, entre os padrões extraídos na etapa de *Data Mining*, os padrões interessantes ao critério estabelecido pelo usuário, podendo voltar à fase inicial para novas iterações. Ao término da avaliação, o conhecimento descoberto deverá ser implantado e incorporado ao sistema, sempre documentando e publicando os métodos, a fim de apresentar o conhecimento descoberto ao usuário. É apresentado na Figura 1 de uma maneira simplificada todas as fases do KDD.

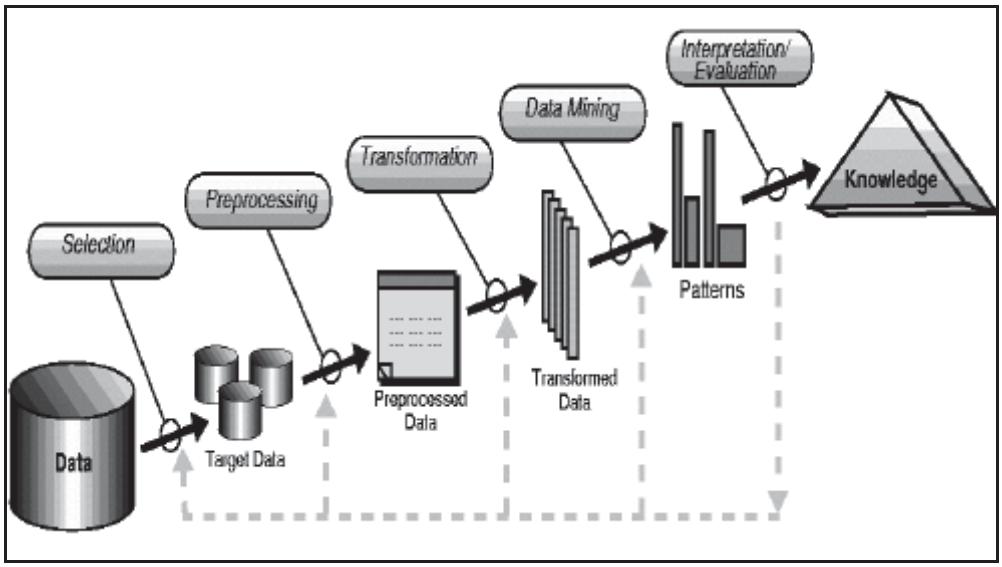


FIGURA 1 – Fases do Processo de KDD. Fonte: FAYAAD *et al.* (1996).

3.2. Mineração de Dados

Data Mining ou Mneração de Dados é o processo de pesquisa em grandes quantidades de dados para extração de conhecimento, utilizando técnicas de Inteligência Computacional para procurar relações de similaridade ou discordância entre dados, com o objetivo de encontrar padrões, irregularidades e regras, com o intuito de transformar dados, aparentemente ocultos, em informações relevantes para a tomada de decisão e/ou avaliação de resultados.

A mineração de dados possibilita a busca em grandes bases de dados de informações desconhecidas, permitindo aos gestores uma maior agilidade nas tomadas de decisões. Uma empresa que utiliza *Data Mining* é capaz de criar parâmetros para entender o comportamento do consumidor, identificando afinidades entre as escolhas de produtos e serviços, prevendo hábitos de compras e analisando comportamentos habituais para detecção de fraudes PINTO (2005).

A principal motivação para a utilização de *Data Mining* é a grande disponibilidade de dados armazenados eletronicamente, com informações úteis e ocultas, podendo auxiliar na previsão de um conhecimento futuro, indo além do

armazenamento explícito de dados. Em outras palavras, descobrir informações sem uma prévia formulação de hipóteses e buscar por algo não intuitivo, é na verdade tornar dados sem obviedade em valiosas informações estratégicas.

3.2.1. Métodos de *Data Mining*

De acordo com FAYYAD *et al.* (1996), existem diversos métodos de *Data Mining* para encontrar respostas ou extrair conhecimento em repositórios de dados, sendo os mais importantes para o KDD: Classificação, Modelos de Relacionamento entre Variáveis, Análise de Agrupamento, Sumarização, Modelo de Dependência, Regras de Associação e Análise de Séries Temporais. Deve-se ressaltar que a maioria desses métodos serve de base para técnicas das áreas de aprendizado de máquina, reconhecimento de padrões e estatística. Um breve resumo dos métodos mais importantes são descritos a seguir.

A Classificação associa ou classifica um item a uma ou várias classes categóricas pré-definidas, utilizando comumente uma técnica estatística chamada análise discriminante, objetivando envolver a descrição gráfica ou algébrica das características diferenciais das observações de várias populações, além da classificação das observações, em uma ou mais classes pré-determinadas.

Os Modelos de Relacionamento entre Variáveis associam um item a uma ou mais variáveis de predição de valores reais, consideradas variáveis independentes ou exploratórias. Regressão linear simples, múltipla e modelos lineares por transformação são as técnicas mais utilizadas para verificar a existência do relacionamento funcional entre duas variáveis quantitativas.

A Análise de Agrupamento, ou *Cluster*, associa um item a uma ou várias classes categóricas (ou *clusters*), determinando as classes pelos dados, independentemente da classificação pré-definida. Os clusters são definidos por meio do agrupamento de dados baseados em medidas de similaridade ou modelos probabilísticos, visando detectar a existência de diferentes grupos dentro de um determinado conjunto de dados e, em caso de sua existência, determinar quais são eles.

A Sumarização determina uma descrição com dispersão reduzida para um dado subconjunto no pré-processamento dos dados, freqüentemente utilizadas na análise de descobrimento de dados. Podemos citar como exemplos simples de sumarização de dados, as medidas de posição e variabilidade.

O Modelo de Dependência descreve dependências entre variáveis. Modelos de dependência existem em dois níveis: estruturado e quantitativo. O nível estruturado especifica, geralmente em forma de gráfico, quais variáveis são localmente dependentes. O nível quantitativo especifica o grau de dependência, usando alguma escala numérica.

As Regras de Associação determinam relações entre campos de um banco de dados, contribuindo para a tomada de decisão, recebendo assim, na Seção 3.3, um maior aprofundamento. Por fim, a Análise de Séries Temporais determinam características seqüenciais, como dados com dependência no tempo. Seu objetivo é modelar o estado do processo extraindo e registrando desvios e tendências no tempo.

3.3. Regras de Associação

O processo de extração de regras de associação foi proposto inicialmente por AGRAWAL *et al.* (1993), e representa um padrão ocorrido em combinações de itens com determinada freqüência em uma base de dados GONÇALVES (2005).

As regras de associação foram definidas através da seguinte formalização AGRAWAL *et al.* (1994): seja $I = \{I_1, I_2, I_3, \ldots, I_n\}$ um conjunto de itens e T um conjunto de transações, sendo cada transação T' um conjunto de itens, de forma que $T' \subseteq I$, e cada T' será associado a um único identificador $T'IT$, tal que $X \subseteq T'$, $X \subseteq I$. Desta maneira a regra de associação gerada terá como forma $X \rightarrow Y$, ou seja, X será o antecedente e Y o conseqüente da regra de associação criada.

Outra abordagem para o conceito de regras de associação leva em consideração os conceitos de Confiança e Suporte GYŐRÖDI *et al.* (2004), em que este estima a probabilidade da ocorrência de um conjunto de itens I em uma transação T', e aquele, para um conjunto Z de itens, refere-se à porcentagem de transações de uma base de dados que contem esses itens de Z.

No Quadro 1 VASCONCELOS *et al.* (2004), tem-se a fórmula utilizada no Suporte, para uma regra $X \rightarrow Y$, onde X e Y são conjuntos de itens, tendo por numerador a quantidade de transações em que ambos, X e Y, estão ocorrendo simultaneamente, e por denominador o número total de transações.

$$Suporte = \frac{Freqüência\ de\ X\ e\ Y}{Total\ de\ T}$$

QUADRO 1 – Suporte.

A Confiança é evidenciada utilizando a mesma regra $X \rightarrow Y$, mas seu denominador representa o valor total de transações em que o item X ocorre VASCONCELOS *et al.* (2004), visto no Quadro 2.

$$Confiança = \frac{Freqüência\ de\ X\ e\ Y}{Freqüência\ de\ X}$$

QUADRO 2 – Confiança.

Na próxima Seção será demonstrado detalhadamente todo o processo de KDD, focando na fase de *Data Mining* e utilizando-se o método descrito nesta.

4. METODOLOGIA

Esta Seção trata da descrição precisa de todos os métodos, materiais, técnicas e equipamentos utilizados para descrever os experimentos realizados, juntamente com os resultados obtidos.

O processo de KDD é evidenciado de fato, com os experimentos utilizando uma base de dados proveniente de uma empresa atuante na área do comércio varejista, optante por um acordo NDA (*Non-Disclosure Agreement*). O *software* de mineração de dados utilizado para realizar a geração de padrões úteis foi o *WEKA* (http://www.cs.waikato.ac.nz/ml/weka/). Desenvolvida na linguagem *Java* pela Universidade de *Waikato* na Nova Zelândia, trabalha com diversas técnicas de *Data Mining*, além de ser um *software* livre e de fácil manuseio.

Todos os resultados desse experimento foram alcançados utilizando-se um computador pessoal (PC) com a seguinte configuração: Intel Core 2 Duo 1,8GHz,

com 2MB de Cache L2, 1GB de memória e 150GB de HD. O sistema operacional utilizado foi o Microsoft Windows XP com Service Pack 2. No momento dos testes, nenhum outra aplicação estava em execução, exceto o WEKA. A Seção 4.1 fará uma abordagem dos experimentos e resultados do estudo de caso proposto.

4.1. Experimentos e Resultados

Como conhecido anteriormente na Seção 3.1, o ponto de partida do KDD encontra-se na escolha ou seleção da massa de dados de acordo com o domínio do objetivo requerido, que neste caso, seria a escolha do perfil do cliente mais rentável para empresa e o grupo de produtos por ele comprado.

Após análise do banco de dados, foi selecionada uma amostra de 102.000 registros identificando as vendas realizadas no período de 01 de janeiro de 2008 a 30 de abril de 2008, de cerca de 10 filiais situadas no estado da Paraíba. Destaca-se na Tabela 1 os atributos necessários ao processo de mineração de dados.

TABELA 1 – Atributos da Base de dados submetidos à mineração.

Atributo	Descrição	Valores Distintos
GRUPO	Grupo no qual o produto foi classificado	4
VALOR	Faixa de valores da compra	6
IDADE	Faixa etária do cliente.	6
SEXO	Sexo do cliente.	2

O atributo GRUPO representa o grupo na qual o produto foi classificado, com 4 valores distintos, a saber: Eletro, Magazine, Móveis, Telecomunicação. Os valores das compras foram decompostos por faixa de valores: abaixo de R$100,00, de R$101,00 à R$300,00, de R$301,00 à R$500,00, de R$501,00 à R$700,00, de R$701,00 à R$900,00 e acima de R$900,00. A idade também foi um atributo organizado em faixa de valores: de 18 a 23 anos, de 24 a 29, de 30 a 37, de 38 a 42, de 43 a 50 e acima de 50 anos. Por fim, o atributo SEXO do cliente.

A fase de *Data Cleaning*, com intuito de eliminar tuplas nulas, valores considerados inconsistentes ou errados, definidos como ruído, e diminuírem redundâncias, reduziu para 87.744 tuplas, sendo posteriormente convertidas para o padrão utilizado no *software WEKA* CORREA (2004). Nesta ferramenta, expôs-se ao processo de mineração de dados, utilizando-se algoritmo de regras de associação *Predictive Apriori* SCHEFFER (2001), GILLMEISTER et al. (2005), obtendo 10 regras de associação, sendo expostas no Quadro 3.

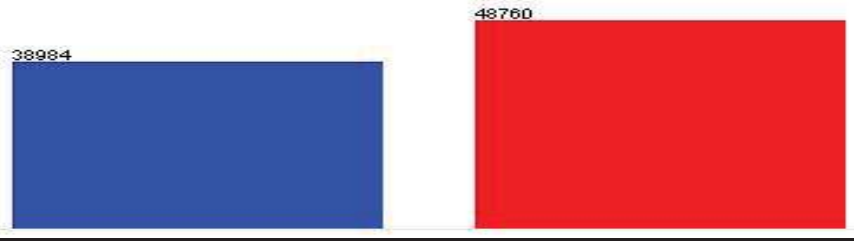
```
Best rules found:

 1. valor=901_MAX idade=18_23 sexo=M 50 ==> grupo=ELETRO 49    acc:(0.96611)
 2. grupo=MAGAZINE idade=18_23 sexo=M 48 ==> valor=0_100 47    acc:(0.9644)
 3. grupo=MAGAZINE valor=301_500 idade=38_42 17 ==> sexo=M 17    acc:(0.96314)
 4. valor=901_MAX idade=18_23 108 ==> grupo=ELETRO 105    acc:(0.96116)
 5. grupo=MAGAZINE valor=301_500 idade=30_37 38 ==> sexo=F 37    acc:(0.95418)
 6. grupo=MAGAZINE idade=24_29 sexo=F 297 ==> valor=0_100 284    acc:(0.95162)
 7. valor=901_MAX idade=51_MAX sexo=M 226 ==> grupo=ELETRO 216    acc:(0.9502)
 8. valor=701_900 idade=30_37 sexo=M 183 ==> grupo=ELETRO 174    acc:(0.94439)
 9. valor=701_900 idade=51_MAX sexo=M 180 ==> grupo=ELETRO 171    acc:(0.94349)
10. grupo=MAGAZINE idade=18_23 278 ==> valor=0_100 263    acc:(0.94284)
```

QUADRO 3 – Regras encontradas.

A ferramenta gera uma boa diversidade de gráficos, utilizando-se das variáveis utilizadas. Em todos os gráficos, o *Label* é o rótulo utilizado para descrever a variável e o *Count* é a contagem desta variável. É descrito no Quadro 4 a distribuição do atributo sexo dos clientes, observando-se a maioria feminina com 48.760 clientes. O sexo masculino contou com 38.984 clientes. Cada atributo, masculino e feminino, será representado pelas cores azul e vermelho, respectivamente, em todos os próximos gráficos.

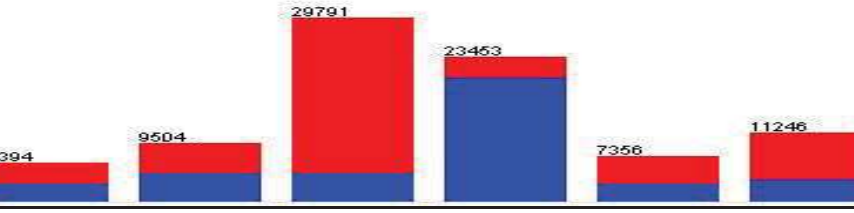
Label	Count
M	38984
F	48760



QUADRO 4 – Atributo sexo dos clientes.

Observa-se no Quadro 5 a distribuição do total de registros de acordo com a faixa de idade, sendo possível verificar que o maior consumo está na faixa dos 30 aos 37 anos, principalmente para o sexo feminino. No segundo maior índice, dos 38 aos 42 anos, esse consumo passa a ser maior para o sexo masculino.

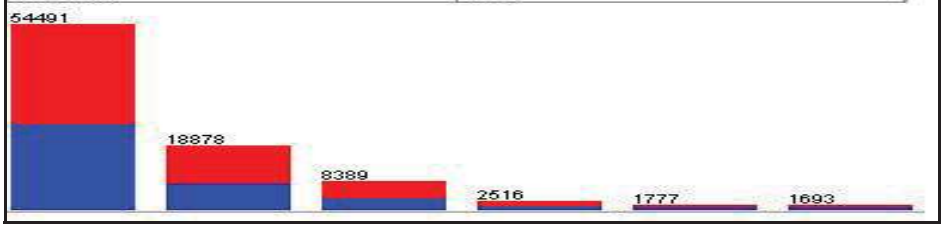
Label	Count
18_23	6394
24_29	9504
30_37	29791
38_42	23453
43_50	7356
51_MAX	11246



QUADRO 5 – Idade dos Clientes.

Com relação ao atributo do valor da compra, é possível observar através do Quadro 6 a incidência maior para compras de produtos que custam até R$100,00, com uma ligeira vantagem do sexo feminino. À medida que a faixa de valores aumenta, nota-se uma grande diminuição da contagem de clientes, o que é normal, visto o nível aquisitivo da maioria dos clientes da empresa.

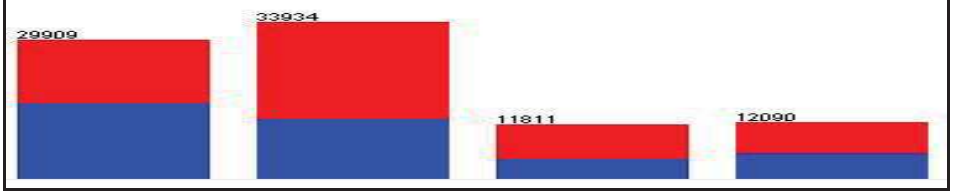
Label	Count
0_100	54491
101_300	18878
301_500	8389
501_700	2516
701_900	1777
901_MAX	1693



QUADRO 6 – Valor da compra.

Os eletros, o que envolve eletro-eletrônicos e eletro-domésticos, são os produtos com maior incidência de compra, sendo seguido de perto pelos produtos da área de magazine. Essa demonstração é possível de verificar observando os dados exibidos no Quadro 7.

Label	Count
MAGAZINE	29909
ELETRO	33934
MOVEIS	11811
TELECOMUNICACAO	12090



QUADRO 7 – Grupo de produtos.

Após o processamento do arquivo pelo *WEKA* e a obtenção dos padrões obtidos na execução da mineração de dados, as regras geradas foram levadas à apreciação dos especialistas de domínio da organização, para análise e avaliação da qualidade das regras obtidas. Segundo os mesmos, quatro regras apresentam um conhecimento novo já que não haviam identificado que no segmento Magazine, com produtos até R$100,00, tinham maior venda entre os indivíduos do sexo feminino até 29 anos, além de verificar a preferência masculina por produtos eletro-eletrônicos de custo acima dos R$700,00. O algoritmo trouxe duas regras com informações duplicadas. As quatro regras restantes foram consideradas óbvias, porém, corretas, levando em consideração o domínio de negócio.

Após a análise das regras e a verificação de um nicho de mercado a ser investido, os gestores optaram por medidas estratégicas específicas, com campanhas de marketing e promoções específicas definidas pelo perfil do cliente.

5. TRABALHOS RELACIONADOS

Mesmo trabalhando em diferentes níveis de detalhamento, algumas pesquisas

encontradas na literatura referem-se ao processo de descoberta de conhecimento em base de dados dando ênfase à fase de mineração de dados e aos algoritmos de *Data Mining*, em especial ao algoritmo *Apriori*, a fim de verificar sua eficiência e eficácia no âmbito da extração de padrões corretos de uma base de dados dando suporte a tomada de decisão em nível de gerencial.

Dentre esses trabalhos que serviram para o embasamento e amadurecimento desta proposta descrita estão: "Utilização de Ferramentas de KDD para Integração de Aprendizagem e Tecnologia em Busca da Gestão Estratégica do Conhecimento na Empresa" realizado por BOENTE et al. (2007). Neste trabalho são abordados os conceitos de Gestão de Conhecimento e de KDD com suas tarefas e métodos, registrando alguns *softwares* utilizados no processo de mineração de dados, sendo necessário adicionar à pesquisa conceitos práticos da efetiva utilização do KDD por meio da empresas.

E os estudos de GILLMEISTER et al. (2005) que demonstram a utilização do algoritmo *Apriori*, e de algoritmos baseados neste, comparando com outros, realizando a geração de regras de associação que trazem um conhecimento novo ao gestor auxiliando assim a tomada de decisão. No entanto, o mesmo apresenta resultados com uma amostra de apenas 7000 registros.

6. CONCLUSÕES E TRABALHOS FUTUROS

O capital humano é um dos maiores responsáveis pelo êxito da GC, seja por um usuário do conhecimento, seja por um fornecedor do conhecimento. A Tecnologia da Informação além de possuir um papel de apoio nos projetos de GC identificando, desenvolvendo e implantando novas tecnologias, contribui para o intercâmbio de conhecimentos e experiências entre as organizações, proporcionando novas perspectivas e soluções, beneficiando um melhor progresso do conhecimento já existente, tornando ágil a tomada de decisão, aperfeiçoando processos, reduzindo custos e aumentando receitas.

Com os resultados obtidos demonstrou-se, na prática, como as diversas tecnologias ligadas ao processo de descoberta de conhecimento em bases de dados podem apoiar as tomadas de decisões, de forma a manter as organizações competitivas com relação à concorrência e, principalmente, manterem-se no mercado.

Como trabalho futuro, pode-se realizar uma comparação entre diversos algoritmos de regras de associação, realizando experimentos que integrem ferramentas de mineração de dados com ferramentas específicas para a GC, apoiando ainda mais o nível gerencial corporativo para aumentar e propiciar melhores condições na tomada de decisões de forma eficaz e eficiente.

REFERÊNCIAS

AGRAWAL, R.; Imielinski T.; Srikant R. "Mining Association Rules between Sets of Items in Large Databases" 1993, Proc. of the ACM SIGMOD Intl. Conf. on Management of Data, Washington, Estados Unidos, p. 207–216;

AGRAWAL, R. et al. "Fast algorithms for mining association rules in large databases" 1994. in: International Conference on Very Large Data Bases, VLDB, 20, Santiago, Proceedings Hove: Morgan Kaufmann, p. 478-499;

ALVARENGA NETO, R. C. D. "Gestão do Conhecimento em Organizações: Proposta de Mapeamento Conceitual Integrativo" 2005. Tese (Doutorado em Ciência da Informação) – PPGCI, Escola de Ciência da Informação da UFMG, Belo Horizonte;

BOENTE, Alfredo Nazareno Pereira *et al.* "Utilização de Ferramentas de KDD para Integração de Aprendizagem e Tecnologia em Busca da Gestão Estratégica do Conhecimento na Empresa" 2007. in: Simpósio de Excelência em Gestão e Tecnologia, Rio de Janeiro, Brasil;

CORREA, A. C. G.; SCHIABEL, Homero "Descoberta de Conhecimento em Base de Imagens Mamográficas" 2004. CBIS'2004 - IX Congresso Brasileiro de Informática em Saúde;

DAVENPORT, T. H., PRUSAK, L. "Conhecimento Empresarial: como as organizações gerenciam o seu capital intelectual" 1998. Campus. Rio de Janeiro.

DAVIS, Gordon B. "Management Information Systems: Conceptual Foundations, Structure and Development". McGraw-Hill. Tokyo;

FAYYAD, U. M., Piatetsky Shapiro, G., Smyth, P. & Uthurusamy, R. "Advances in Knowledge Discovery and Data Mining" 1996, AAAIPress, The MitPress;

GILLMEISTER, Paulo Ricardo Guglieri; Cazella, Sílvio César (2005) "Uma Análise Comparativa de Algoritmos de Regras de Associação: Minerando Dados da Indústria Automotiva", Centro de Ciências Exatas e Tecnológicas, UNISINOS, São Leopoldo, RS;

GONÇALVES, Eduardo Corrêa "Regras de Associação e suas Medidas de Interesse Objetivas e Subjetivas", UFF - Universidade Federal Fluminense, in INFOCOMP – Journal of Computer Science, Rio de Janeiro, RJ;

GYŐRÖDI, Cornelia et al. "A Comparative Study of Association Rules Mining Algorithms", in: 1st Romanian, Hungarian Joint Symposium on Applied Computational Intelligence, Oradea, Romania;

KRUGLIANSKAS, Isak e TERRA, José Cláudio Cyrineu "Gestão do Conhecimento em Pequenas e Médias Empresas". Rio de Janeiro: Campus, 2ª. Edição;

PINTO, F.; Santos M. F. "Descoberta de Conhecimento em Bases de Dados em Actividades de CRM"; Datagadgets 2005; 1º Congresso Espanhol de Informática CEDI 2005; Granada;

SANCHES, Osvaldo Maldonado "Planejamento Estratégico de Sistemas de Informação Gerencial", REVISTA DE ADMINISTRAÇÃO PÚBLICA (RAP) de jul./ago de 1997, Fundação Getúlio Vargas (FGV). Rio de Janeiro, RJ

VASCONCELOS, L. M. R. de; Carvalho, C. L. de (2004) "Aplicação de Regras de Associação para Mineração de Dados" na Web. UFG. Relatório Técnico. Goiás, GO;

WU, J. H.; WANG, Y. M. "Measuring KMS success: A respecification of the DeLone and McLean's model". Information & Management, v.43, n.6, p.728., setembro, 2006. Retrieved December 7, 2006, from ABI/INFORM Global database. (Document ID: 1150986541).