



## DATA MINING NO CONTEXTO DE CUSTOMER RELATIONSHIP MANAGEMENT EM UMA FRANQUIA COCA – COLA COMPANY

Renata Azevedo Santos Carvalho<sup>1</sup>

### RESUMO

Data Mining é uma área de pesquisa multidisciplinar, incluindo tecnologias de banco de dados, inteligência artificial, redes neural, aprendizado de máquina, estatística e visualização de dados, tendo como objetivo específico a descoberta de conhecimento novo que por ventura esteja escondido em grandes massas de dados. Como um dos grandes objetivos de uma corporação é conhecer seus clientes, este conhecimento precisa ocorrer em vários níveis, desde o tipo de produto desejado até que tipo de ofertas os clientes estão dispostos a aceitar mesmo que os produtos não sejam essenciais no momento. Esta forma de mercado dirigido pode atingir o extremo de uma relação individual com cada cliente à medida que a empresa deseje investir em segmentações (classificações) sucessivas da sua clientela. Sendo assim, esse trabalho tem como finalidade aplicar técnicas de mineração em conjunto com as diretrizes do CRM à uma franquia da Coca-Cola afim de gerar uma nova classificação dos seus clientes e auxiliar o cumprimento das metas anuais de venda com a criação de novas atividades de marketing dado o resultado da análise dos dados minerados.

**Palavras-chave:** Data Mining; descoberta de conhecimento; CRM; KDD

### Introdução

O mundo passou por profundas e importantes transformações, sobretudo nos últimos cinco anos, impulsionadas pelo crescimento da Internet. Ao ganhar na Web mais um poderoso canal de comercialização e de comunicação, o setor corporativo precisou rever alguns conceitos e se reestruturar. Na era digital, tudo é muito rápido. Para poder competir nesse cenário, não basta oferecer produtos e serviços com melhor qualidade e preço. É preciso, também, conhecer o consumidor, ser capaz de satisfazê-lo e não perdê-lo, logo a

---

<sup>1</sup> Mestre em Ciência da Computação pela Universidade Federal de Pernambuco – UFPE. MCTS em Sql Server 2008. Professora universitária da Faculdade de Administração e Negócios de Sergipe – FANESE. Analista de Banco de Dados do Sergipe Parque Tecnológico - SERGIPETEC. renatamsdn@gmail.com



seguir, para a concorrência. Em outras palavras, se faz necessário saber criar e gerenciar o relacionamento com o cliente de forma a gerar valor para a companhia.

Não é por acaso que um dos temas que mais tem chamado a atenção das empresas e que vem sendo amplamente debatido em eventos específicos e na mídia especializada é justamente o CRM – *Customer Relationship Management*[43]. Uma filosofia que envolve pessoas, processos, tecnologia, visando a criação de uma sistemática para adquirir maior conhecimento sobre o cliente ao longo de toda a vida dele, e não apenas no momento em que realiza uma transação comercial com a empresa.

Paralelamente a isto, nos deparamos também com o crescimento da utilização de sistemas informatizados para o armazenamento e gerenciamento de informações sobre os clientes e suas transações com as empresas prestadoras de serviços. Este relacionamento gera uma enorme quantidade de informação guardada nos bancos de dados, informação esta que é potencialmente importante, mas que ainda não foi descoberta. Estas informações podem ser importantes para a empresa conhecer melhor e entender seus clientes oferecendo assim produtos voltados às suas necessidades.

Os dados dos sistemas transacionais, os quais são a fonte para o CRM operacional, aliados às necessidades dos clientes, obtidas através das centrais de relacionamento (CRM analítico), podem gerar uma importante fonte de informações. Porém, nem sempre isso acontece, pois os dados são armazenados separadamente no repositório de cada departamento, impossibilitando a integração dos dados para que se transformem na inteligência do negócio.

Neste momento, entra o processo conhecido como *data mining* [29] que desenvolve um processo organizado de transformação de dados em conhecimento, incorporando ferramentas de tecnologia de informação, conhecimento de gestão da informação e análises de dados sofisticadas, executando assim, do ponto de vista estatístico, uma análise exploratória multivariada de grandes quantidades de dados numéricos.

A empresa escolhida no estudo de caso possui um vasto volume de dados armazenados, trata-se de uma franquia da *Coca-Cola Company* e atua nos mercados de Alagoas, Sergipe e Bahia utilizando sistemas informatizados tanto para a venda quanto para a coleta de pesquisas de opinião. Esses dados, porém, não são aproveitados para a geração de informação de forma eficiente. Observa-se também um sistema completamente engessado de classificação de seus clientes e pouco conhecimento sobre o comportamento dos mesmos.



Neste trabalho busca-se a solução de um problema de gerenciamento de relação com o cliente através da criação, estudo e otimização do processo de Mineração de Dados como um todo. A questão de fidelização de clientes a uma determinada indústria produtora e distribuidora de bebidas no Brasil, bem como um possível aumento das vendas de produtos será avaliada em busca de uma solução satisfatória. Um processo de data mining cuidadoso, partindo da análise do problema, passando pelo tratamento dos dados e culminando na busca e otimização do modelo que mais se adéqüe a tarefa será a arma para solucionar a questão do aumento da fidelização dos clientes. Objetiva-se também demonstrar o grande valor e vantagem competitiva que podem ser gerados com a utilização da Mineração de Dados em estratégias de retenção de clientes, através da análise da lucratividade que o sistema de retenção de clientes desenvolvido irá gerar para a empresa.

## Metodologia

O estudo utiliza com referências manuais e materiais sobre o CRM e o Data Mining que estão disponibilizados em diversas bibliografias científicas, sendo objeto de análise criterioso. Também foi estudada a ferramenta *Weka* (*Waikato Environment for Knowledge Analysis*)

Inicialmente estudou-se como é o processo de instalação da plataforma, em seguida como seriam feitas as adaptações ao *Weka* para que na etapa futura da pesquisa possa ser estudado o seu impacto no que se refere a obtenção no ganho de lucro com o aumento da fidelização dos clientes da empresa onde o estudo de caso está sendo executado.

## Resultados e Discussão

Foi feita uma análise dos resultados da classificação e da associação sob a ótica do negócio pelo especialista de *marketing*, a seguir alguns comentários das 3 regras expostas abaixo, referentes ao algoritmo *A priori*, nela os resultados foram vistos pela ótica do minerador.

5. COCA-COLA_K5=n FANTA_K5=n KUAT_N5=n 36839 ==> KUAT_K5=n 36680 conf:(1)
6. COCA-COLA_K5=n COCA-COLA_N5=n FANTA_K5=n 37486 ==> KUAT_K5=n 37321 conf:(1)
49. COCA-COLA_LATA=y 4739 ==> FANTA_N5=n 4300 conf:(0.96)

1. Resultados obtidos com as regras 5, 6 e 49



Com relação aos resultados da classificação e a forma como ela foi gerada com suas condições, ele achou bastante interessante por levar em consideração fatores de situação financeira e de pendências tanto com a empresa como com o mercado externo. Esse parâmetro nunca tinha sido levado em conta antes.

Outra característica que também foi de grande contribuição foi a utilização da característica lucro que informa o *payback* daquele cliente com a empresa, ou seja, o quanto do investimento feito nele já obteve retorno.

Já na associação ele comentou os resultados obtidos, nas regras expostas, sendo assim:

- ✓ A regra (5) o especialista após análise das atividades praticadas detectou falha em algumas delas e concordou com o resultado dessa mineração e na descoberta desses erros.
- ✓ A regra (49) gerou a descoberta de um padrão que pode implicar em um perfil bastante crítico de pessoas que ao comprarem um determinado refrigerante se fidelizam a ele, não comprando mais nenhum da franquia. Atividades como criação de embalagens mix com mistura de produtos poderia ajudar a solucionar esse problema, pois o consumidor passaria a experimentar novos refrigerantes e esse padrão acabaria sendo minimizado.
- ✓ A regra (6) sugere com bastante confiança uma inconsistência, pois informa novamente um conjunto de produtos que a sua compra conjunta não acontece, implicando com isso a não compra de um produto Kwat.

A partir de uma análise geral do especialista do domínio, puderam-se obter as seguintes impressões:

- ✓ A compra da marca Coca Cola é predominante e acaba excluindo a compra de outras marcas;
- ✓ Detecção de atividades de *marketing* feitas erroneamente, prejudicando com isso as vendas de algumas marcas de refrigerantes;
- ✓ Exclusão de qualquer produto quando o cliente resolve comprar a Coca-Cola em lata;
- ✓ Nem sempre a compra de um item tem boa correlação com os demais produtos. Isso pôde ser comprovado na maioria das associações.

Para os resultados do algoritmo *PredictiveApriori*, o especialista viu possíveis ações a serem tomadas:

- ✓ Criação de promoções para a compra de Fanta e Kwat juntas;



- ✓ Formulação de novos preços de venda para esses clientes do produto Fanta\_NS;
- ✓ Geração de embalagens mix com produtos, por exemplo, Coca\_ks e Fanta\_ks ou Coca\_ns e Kuat\_ns, forçando a compra de mais de uma marca por vez.

Após a tarefa de classificação conseguimos obter quantidades de clientes para cada quadrante de classificação da empresa, como o estudo focou nos clientes Líder Alvo, imagina-se o seguinte possível cenário: a empresa possui trinta mil clientes, destes cinco mil são clientes alvos, utilizando uma janela de dois meses para a atividade de *marketing* fazer efeito, um custo de incentivo de R\$3.000,00 ao longo desse tempo e o lucro já conhecido de que um cliente fiel gera para a empresa algo em torno de R\$2.000,00 por mês, pode-se dizer que em quatro meses a empresa já terá um lucro de R\$1.000,00 só com fidelização de um cliente. Caso o sucesso das atividades de *marketing* sobre esses clientes eleitos para a fidelização seja algo em torno de 5% significará um lucro para a empresa de R\$250.000,00 em quatro meses. Essas informações financeiras foram obtidas através de consulta ao departamento de *marketing* da empresa em questão e repassadas pelo gerente de vendas.

É importante concluir que o sistema de auxílio à fidelização proposto deve ser empregado em clientes de alto valor e por isso ele foi focado nos clientes Líder Alvo.

Os resultados obtidos demonstram a aplicabilidade do sistema e os ganhos financeiros que ele pode prover para empresas que decidam utilizá-lo como parte de suas estratégias de fidelização de clientes. Diante de estudos observou-se que se uma empresa aumenta em 5% a quantidade de clientes fiéis elas podem ter um incremento nos lucros de 25 a 85%, e que caso o sistema fosse empregado, esse aumento poderia ser alcançado facilmente. De qualquer forma, o sistema não é totalmente automático e foi dada muita ênfase ao fato de que a compreensão do negócio é essencial em quase todas as etapas do sistema. Sem esse entendimento o sistema perde muito do seu poder de auxílio e aderência ao negócio.

## Conclusões

Neste trabalho foi detalhado um processo de Mineração de Dados, primeiramente diferenciando mineração em si do processo de KDD, o que não raro é confundido na literatura. Depois foi realizado um estudo de métodos e técnicas, tanto para a preparação de dados quanto para a modelagem dos mesmos. Foram apresentadas o conceito sobre o



processo de CRM e o marketing direto, formas mais modernas de se aplicar marketing com os clientes, sendo que a sua grande contribuição foi justamente a definição de um processo e um modelo, embasado em tecnologias existentes, para que empresas usem a massa de dados que possuem sobre seus clientes, para entendê-los e, com isso, tratá-los de maneira diferenciada, que é justamente a proposta do CRM, mostrando com isso, como integrar duas áreas de ação ou duas áreas de conhecimento (KDD e CRM), tirando o melhor proveito possível delas.

Com relação ao estudo de caso, este trabalho cumpre com o objetivo geral relacionado à criação de uma nova classificação para a empresa em estudo bem como uma análise dos perfis de compra dos seus clientes gerando novas atividades de marketing já no ambiente de produção da mesma. Com isso, pôde-se observar também a confiança da alta gerência em apoiar essa nova ferramenta, que eles avaliaram como um grande diferencial para as demais franquias.

## Referências

- [1] FAYYAD, Usama; PIATETSKY-SHAPIRO, Gregory; SMYTH Padhraic; UTHURUSAMY, Ramasamy -Advances in Knowledge Discovery and Data Mining. AAAI/MIT Press, 1996.
- [2] BERRY, Michael; LINOFF, Gordon -Mastering Data Mining: The Art and Science of Customer Relationship Management - John Wiley & Sons, 2000.
- [3] BOZDOGAN, Hamparsum -Statistical Data Mining & Knowledge Discovery – Chapman & Hall/CRC, 2003.
- [4] KLOSGEN, Willi; ZYTKOW, Jan M.; ZYT, Jan -Handbook of Data Mining and Knowledge Discovery - Oxford University Press, 2002.
- [5] PYLE, Dorian -Data Preparation for Data Mining - Morgan Kaufmann, 1999. [6] INMON, William H.; ZACHMAN, John A -Data Stores Data Warehouse and the Zachman Framework – Managing Enterprise Knowledge - Editora Campus, 1996.
- [7] FÉLIX, Mödritscher -Data Mining No Processo De Extração De Conhecimento De Bases De Dados -Dissertação De Mestrado, ICMC-USP, 1998.
- [8] GLYMOUR, Clark; MADIGAN, David; PREGIBON, Daryl; SMYTH, Padhraik - Statistical Themes and Lessons for Data Mining - Data Min. Knowl. Discov. 1(1): 11-28, 1997.



[9]GROTH, Robert - Data mining - Prentice Hall, Inc. 1998.

[10]LOBO, Alexandre -Marketing de relacionamento - São Paulo: Seminários SSJ, 2002.

[11]SWIFT, Ronald - CRM: Customer Relationship Management, o revolucionário marketing de relacionamento com o cliente - Rio de Janeiro: Prentice Hall, 2001. p. 3-4, 12, 331.

[12]GREENBERG, Paul - CRM, Customer Relationship Management na velocidade da luz: conquista e lealdade de clientes em tempo real na Internet - Rio de Janeiro: Campus, 2001. p.60, 389-392.

[13]PEPPERS, William;ROGERS, Michael;DORF, Bob - One to One Field Book - N.Y: Currency Book, 1999.