

## UTILIZAÇÃO DO DATA MINING COM ÊNFASE NO COMPORTAMENTO DO CONSUMIDOR: UM ESTUDO DE CASO EM UM E-COMMERCE

Alisson Rodrigues Vitorino<sup>1</sup>

Homero Domingues<sup>2</sup>

### RESUMO

Ferramentas como Data Mining estão cada vez mais presentes para auxiliar na tomada de decisão das empresas, gerando insumos aos gestores, agregando valor aos dados gerados. Este artigo propõe a aplicação da técnica de Data Mining demonstrando sua eficiência em descobertas de conhecimento útil. O objetivo é gerar informações úteis para a tomada de decisão em marketing em uma empresa de e-commerce de vinho, utilizando-se de um modelo não usual. Os resultados obtidos podem auxiliar na definição de estratégias adotadas pelo marketing da empresa. Conclui-se que para o problema proposto, o modelo atende às expectativas de descoberta de padrões e comprovaram que essa técnica pode ser utilizada para auxílio na tomada de decisão desse tipo de empresa.

**Palavras-chave:** Data Mining. E-commerce. Comportamento Consumidor.

### ABSTRACT

Tools such as Data Mining are increasingly present to aid in decision making of enterprises, generating inputs to managers, adding value to data generated. This article proposes the application of Data Mining technique showing its efficiency in discoveries of useful knowledge. The purpose was to do useful information for marketing decision-making in a wine e-commerce company. An unusual model was used, which generated results that can help in the definition of strategy adopted by the marketing of the company. It is concluded that for the problem proposed, the model meets the expectations of finding patterns and proved that this technique can be used to aid in the decision-making process of this type of company.

---

<sup>1</sup> Pós-graduando do Programa de Pós-graduação *lato sensu* do IGTI (Instituto de Gestão e Tecnologia da Informação) em Inteligência de Negócios.

<sup>2</sup> Administrador, Mestre em Administração, Professor e Coordenador de Cursos da Faculdade Doctum de João Monlevade e Orientador de Cursos de Pós-graduação *lato sensu* do IGTI (*Instituto de Gestão e Tecnologia da Informação*, [homerodomingues@yahoo.com.br](mailto:homerodomingues@yahoo.com.br))

**Keywords:** Data Mining. E-commerce. Customer Behavior.

## INTRODUÇÃO

As empresas têm buscado cada vez mais utilizar informações provenientes dos seus dados gerados para fomentar suas estratégias. Pois, na atual era globalizada em que se vive, onde é cada vez maior a competitividade entre as empresas, informação e conhecimento são elementos fundamentais para obter diferenciais mercadológicos frente à concorrência.

Nos últimos anos foi dada uma grande importância para o consumidor, e a partir disso, as empresas tentam oferecer o melhor produto/serviço para o cliente, com intuito de torná-los fiéis. Isso porque “Os clientes estão exigindo cada vez mais qualidade e serviço superiores, além de alguma customização. Eles percebem menos diferenças reais entre produtos e mostram menos fidelidade a marcas” (KOTLER, 2000, p.48).

Porém, entender o comportamento do consumidor não é uma tarefa simples, eles estão mais difíceis de agradar, mais inteligentes em relação ao comportamento de preço, pois possuem mais acesso as informações do mercado. É uma nova era de concorrência, “com o advento da internet, e da globalização desintegrando as fronteiras, as empresas já não podem ficar confiantes em suas participações de mercado e em suas posições competitivas” (PETER, 2005, p.31).

“Hoje em dia, os clientes estão tomando mais iniciativas para determinar o que e como comprar. Eles navegam pela Internet, procuram informações e avaliação dos produtos ou serviços oferecidos, conversam com fornecedores e usuários e decidem quanto à melhor oferta.” (KOTLER, 2000, p.282)

Entender o comportamento do consumidor e conhecer os clientes não são tarefas simples. O comportamento de compra é influenciado por fatores culturais, sociais, pessoais e psicológicos, segundo Kotler (2000). Entretanto, conforme Batista e Fraga (2016), nas últimas décadas, com a tecnologia disponível, as organizações construíram banco de dados com diversos elementos sobre seus clientes e

comportamentos, proporcionando assim subsídio para gestão de relacionamento com o cliente.

Desta forma, atualmente existe um grande armazenamento de informações no banco de dados nas empresas; entretanto, esses dados isolados geram apenas análises superficiais. Para Madruga (2006), a tecnologia vem auxiliando as empresas a desenvolver e trabalhar o marketing de uma forma mais eficiente, utilizando informações de clientes para o seu próprio benefício. Nesse contexto, surge o Data Mining, “uma ferramenta para explorar esse grande volume de dados, descobrindo padrões e relações nessas informações, realizando descobertas importantes” (CARDOSO, 2008, p. 497)

Além disso, o processo de busca por padrão e relações pode ser automatizado, desenvolvendo aplicações que venham extrair dos bancos de dados informações críticas, objetivando assim, estreitar ainda mais o relacionamento da empresa com os consumidores. “A utilização adequada da mineração de dados permite que as empresas transformem o novo conhecimento em ação e valor de negócios” (KUMAR e REINARTZ, 2005, p.235).

As aplicações são diversas, “em bancos, a ferramenta de data mining já vem sendo utilizada desde 1998, com os resultados obtidos reduzindo os custos de venda de produtos de 15% a 30%, além do ganho de produtividade. ” (DA COSTA et. al.). E, segundo Queiroga (2005), sua utilização também pode ser vista em casos de reconhecimento de fraudes em áreas de telefonia ou de cartão de crédito, baseado em perfis de consumidores e também em seus padrões de consumo,

Logo, a escolha deste tema visa a uma contribuição para fomentar a mineração de dados no contexto da inteligência de negócios como uma ferramenta de apoio à tomada de decisão para ser utilizada em maior escala, principalmente, no planejamento estratégico e ações de marketing.

O presente trabalho aplicou técnicas de *Data Mining* com objetivo de gerar informações úteis para a tomada de decisão em marketing em uma empresa de e-commerce de vinho. E como parte do processo de descoberta de conhecimento em banco de dados, este artigo tem como objetivos específicos: testar um modelo

probabilístico de frequência de compra utilizando uma técnica de *Data Mining* e, portanto, sugerir ações baseadas na análise do resultado deste processo.

Este trabalho está dividido em seis sessões, já foram expostos, o tema, objetivo e justificativa. Na sessão dois, é apresentada a metodologia utilizada, na sessão três, dando um embasamento teórico são apresentados os principais conceitos relacionados ao *data mining* e comportamento de compra. Em seguida é feita a apresentação da pesquisa com o estudo de caso, descrevendo um pouco do modelo proposto e os benefícios gerados com este em relação ao comportamento do cliente e ao final serão apresentados a conclusão e as referências.

## 1. METODOLOGIA

Quanto aos objetivos, para este artigo foi utilizada a pesquisa exploratória, com o procedimento de levantamento bibliográfico e documental das principais referências relacionadas aos termos e conceitos de *Data Mining* aplicado, assim como no auxílio do planejamento estratégico do marketing que forem pertinentes ao desenvolvimento do artigo.

A pesquisa exploratória por ser um tipo muito específica, quase sempre ela assume a forma de um estudo de caso Gil (2008). Por isso, neste trabalho apresentou-se uma pesquisa classificada como um estudo de caso, tendo como objeto uma empresa de *e-commerce* de vinhos do país. Pois, segundo Yin (2005), um estudo de caso é uma investigação empírica onde o pesquisador tem pouco controle sobre os acontecimentos, quando o foco se encontra em fenômenos contemporâneos inseridos em algum contexto da vida real, que busca fontes de evidências para responder ao problema de pesquisa, e que se beneficia do desenvolvimento prévio de proposições teóricas que conduzirão a coleta e análise de dados.

Esta pesquisa envolve um estudo intenso para tentar inferir o comportamento de compra do cliente através de elementos gerados pelo próprio cliente, considerando que no banco de dados da empresa é possível encontrar informações comportamentais dos clientes, tais como: frequência e recência de compra, ticket médio, entre outras que

podem ser utilizadas pela *Data Mining*, no intuito de prever o comportamento desses consumidores.

Para alcançar os objetivos é necessário optar por uma abordagem quantitativa. Uma vez que a análise de dados segundo Yin (2005), “consiste em examinar, categorizar, classificar em tabelas, testar ou, do contrário, recombinar as evidências quantitativas e qualitativas para tratar as proposições iniciais de um estudo”. Neste sentido, utilizou-se como procedimento a coleta de informações relativas a toda a base de clientes pertencentes a um e-commerce de vinho do Brasil, que gerou informações transacionais dos clientes ao longo de 4 anos, capturadas e salvas no banco de dados da organização.

Para isso, utilizou-se informações relativas a uma amostra da base de clientes pertencentes a este e-commerce, que gerou informações transacionais dos consumidores ao longo de cinco anos (2012 a 2016), capturadas e salvas no banco de dados da empresa.

## **2. REVISÃO DE LITERATURA**

### **2.1. Data Mining**

Esta seção apresentará características do processo de *Knowledge Discovery in Database*<sup>3</sup> (KDD), tendo como foco a principal etapa deste processo, o *Data Mining*, pontuando assim suas fundamentais tarefas e técnicas.

Com base no aumento sem precedentes da geração de dados nas variadas aéreas do mercado, há uma crescente necessidade - do ponto de vista do negócio - de extrair informações úteis a partir destes dados. O processo de identificar novos conhecimentos desse vasto volume de dados é o objetivo do *Data Mining*.

Sob a perspectiva de Queiroga (2005), o termo *Data Mining* é utilizado para caracterizar uma das etapas de descoberta de conhecimento em bancos de dados KDD. Esses dois termos foram conceituados e distinguidos por Fayyad (et. al., 1996),

---

<sup>3</sup> KDD – Descoberta de conhecimento em base de dados, é o processo de transformação de dados em conhecimento.

no qual ele se refere ao KDD como processo geral de identificar informações úteis dos dados, e o *Data Mining*, a aplicação específica de algoritmos para extração de padrões nos dados, segundo a citação de MOREAS SOUZA E FIGUEIREDO (2014).

Este processo é constituído nas seguintes etapas, conforme a Figura 1.

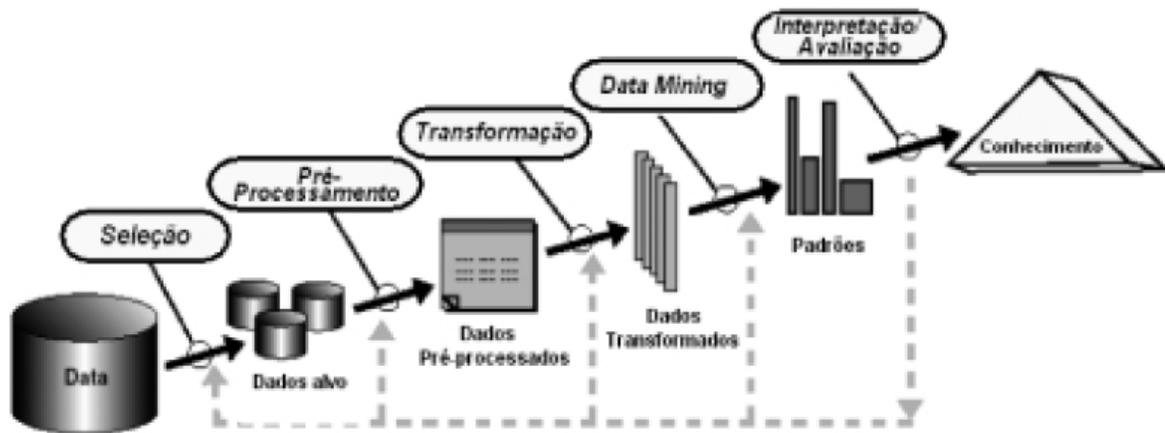


Figura 1 Visão geral das etapas que constituem o processo KDD.

Fonte: Fayyad et al., 1996b

Sandri (2009) faz uma citação sobre cada etapa do KDD, sob a perspectiva de Figueira (1998):

- Seleção: Etapa de organização e agrupamento, organização dos dados;
- Pré-processamento: Esta é uma etapa de adequação dos dados, para que estejam no formato correto e não apresentem falhas, como duplicidade;
- Transformação: Esta é uma fase facilitadora do processo de *Data Mining*;
- *Data Mining*: Principal etapa do KDD é o momento de aplicações de algoritmos para descoberta de padrões;
- Interpretação e avaliação: É o momento de interpretação e validação das novas informações geradas, verificando se respondem ao problema proposto.

Dando ênfase na etapa de Mineração de Dados ou *Data Mining*, ela é a área específica do KDD que trata das técnicas e algoritmos utilizados na detecção dos padrões de dados, é possível dividi-la em algumas etapas:

- Seleção da tarefa;

- Escolha do algoritmo;
- Mineração;
- Descoberta de padrões
- Interpretação do conhecimento descoberto

Após pesquisa bibliografia percebe-se que a definição mais citada e dada como a principal a respeito do *Data Mining* foi criada por Fayyad (et. al,1996, p.4) e “...o processo não-trivial de identificar, em dados, padrões válidos, novos, potencialmente úteis e ultimamente compreensíveis” (CARDOSO E MACHADO, 2008).

Elmasri e Navathe (2005) se referem à mineração de dados como o descobrimento de novas informações em função de padrões em amplas quantidades de dados. Outros autores também tiveram grande participação no estudo de *Data Mining* e definições semelhantes, como, Martinhago (2005, p.45) ao dizer que este processo se trata da busca de padrões, classificações, ou seja, descobrir novas relações, não identificadas a “olho nu”.

De acordo com Da Silva, Boscarioli e Peres (2003) o *Data Mining* é o ramo da computação que utiliza ferramentas e técnicas computacionais com a finalidade de sistematizar o processo de extração de conhecimento útil de grandes volumes de dados”.

De forma geral, o processo de *Data Mining* permite realizar valiosas descobertas em vasto volume de dados gerados através de técnicas de identificação e reconhecimento de padrões. E o novo conhecimento gerado é que faz com que este processo seja considerando uma ferramenta valiosa à tomada de decisões para os negócios. Principalmente em estratégias e ações de marketing fundamentadas em padrões de consumo dos clientes (QUEIROGA, 2005).

## **2.2. Comportamento do Consumidor**

O comportamento de compra complexo envolve um processo de três etapas. Primeiro, o comprador desenvolve crenças sobre o produto. Segundo, ele desenvolve atitudes sobre o produto. Terceiro, ele faz uma escolha refletida. ” (KOTLER, 2000,

p.199).

Já Engel, Blackwell e Miniard (2000), define o comportamento do consumidor em atividades diretamente envolvidas na obtenção, consumo e descarte do produto e serviços e divide este em quatro etapas: primeiro, com os estímulos inputs; segundo, passa para o processamento de informações; na sequência, o processo decisório; e na última etapa constam as variáveis que influenciam o processo decisório.

O próprio Kotler (2000) também diz que o comportamento de compra é influenciado por fatores culturais, sociais, pessoais e psicológicos. Outro autor também cita esses fatores no comportamento de compra do consumidor, Rocha (2004), diz que além dos fatores culturais, sociais, pessoais e psicológicos, este também sofre influência de fatores de natureza política, econômica, tecnológica, ambiental, assim como, fatores mercadológicos como o produto, o seu preço, a sua promoção e o seu ponto de distribuição.

“O comprador pesquisará o que está disponível no mercado, mas acabará comprando com relativa rapidez, devido talvez a um bom preço ou à conveniência da compra.” (KOTLER, 2000, p.199). Após comprar o produto, o consumidor experimenta algum nível de satisfação ou insatisfação.

Kotler (2000) salienta que a satisfação do cliente surge da proximidade entre as expectativas do comprador e desempenho percebido do produto. Se o segundo, não alcançar totalmente as expectativas, o cliente fica desapontado; se alcançar as expectativas, ele fica satisfeito, e se exceder as expectativas ele fica encantado. Essas emoções determinam se o cliente voltará a comprar o produto ou na mesmo na própria loja.

### **3. APRESENTAÇÃO DA PESQUISA**

Este trabalho centra-se em um estudo de caso em um comércio eletrônico de vinhos no Brasil. O e-commerce analisado neste artigo tem a possibilidade de obter informações mais ricas, que auxiliem no processo de tomada de decisão e ações mais eficientes de marketing, direcionando os investimentos de acordo com o perfil do seu cliente.

Nos bancos de dados da empresa é possível encontrar informações comportamentais dos clientes, tais como, frequência e recência de compra, ticket médio, dentre outras, que podem ser utilizadas pelo *Data Mining* no intuito de extrair conhecimento mais útil, no caso deste estudo, prever o comportamento de compra dos consumidores.

Dentro do processo de *Data Mining* existem vários métodos de abordagem entre os mais conhecidos estão Indução de Regras, Redes Neurais e Aprendizado Baseado em Casos. Porém, este estudo pretende usar uma combinação de variáveis geradas pelas transações dos clientes, juntamente com modelo probabilístico conhecido como BG/BB (beta-geométrico / beta-binomial), desenvolvido por Fader, Hardie & Berger (2004) para prever e prever comportamentos dos consumidores de uma empresa de um e-commerce de vinho do Brasil.

Em relação às aplicações de técnicas de mineração de dados, atualmente é facilitada devido à existência de ferramentas que dispõem de recursos de análise de dados e implementam algoritmos específicos. Contudo, esta pesquisa utilizou uma técnica não usual de *Data Mining*, por isso optou-se pela ferramenta  $R^4$ , que é um software livre e muito bem-conceituado no meio acadêmico e profissional. Sendo que mais adiante foi necessário usar o pacote BYTD<sup>5</sup>, já que este possui as funções do modelo BG-BB.

Desta forma, busca-se elaborar uma classificação dos clientes deste e-commerce, que como resultado seja possível prever o tempo residual e o cálculo do *Lifetime Value (LTV)* dos clientes. Segundo Switt (2001) este segundo indicador é um dos mais úteis extraídos dos dados de uma empresa. Os valores obtidos podem auxiliar na definição de estratégia adotada em cada grupo ou até mesmo em ações personalizadas, para os clientes de toda a companhia, pois com eles é possível identificar os consumidores de maior rentabilidade e os mais leais.

Neste trabalho não se pretendeu demonstrar as aplicabilidades de todas as etapas do KDD, apesar de que todos os passos foram aplicados com seriedade e empenho, mas – *como já mencionado* - o foco do estudo é apenas na etapa do *Data*

---

<sup>4</sup> R é um software estatístico com grande poder computacional. <https://www.r-project.org/>

<sup>5</sup> Pacote computacional do R com funções estatísticas.

*Mining*, com um modelo não usual neste tipo de metodologia. Outro ponto que deve ser comentado é que, não houve dificuldade em se trabalhar com os dados da empresa, pois não há exposição de nenhuma informação sigilosa ou de caráter pessoal dos clientes.

### 3.1. Modelo BG-BB

De acordo com De Araujo Ribeiro e Sant'Anna (2007), os modelos de análise de base de clientes buscam prever comportamentos futuros de compras dos consumidores, baseado no histórico transacional, identificando quais são mais propensos a estarem ativos ou inativos em certo período.

O modelo de estudo aqui, BG-BB (beta-geométrico / beta-binomial), foi desenvolvido por Fader, Hardie & Berger (2004) e é utilizado onde o tempo de ocorrência de compras pode vir a ser tratado como uma variável discreta. Para sua aplicação, apenas três informações sobre o comportamento de compra do cliente são suficientes:

1 - A quantidade de períodos (neste estudo de caso, utilizou-se meses) que se realizou as compras;

2 - O último período que ele realizou uma compra;

3 - O total de períodos que o cliente poderia ter realizado compras.

Uma vez obtido essas informações, é possível gerar algumas medidas importantes:

I. A probabilidade do consumidor estar ativo em um determinado período;

II. A quantidade estimada de períodos em que ele estará ativo em um determinado horizonte;

III. O número descontado esperado de transações que ele terá até final do seu relacionamento com a empresa (*discounted expected transactions* – DET). O DET multiplicado pelo lucro médio mensal do cliente gerará uma estimativa para o valor do tempo de vida do cliente (CLV).

Os autores De Araujó Ribeiro e Sant'Anna (2007) salientam que o maior interesse no emprego do modelo BG/BB está no cálculo do *Lifetime Value* do cliente (CLV), ou valor do tempo de vida do cliente, pois o através do modelo empregado, é possível levar em consideração a flutuação nas futuras transações (descontadas para valor presente).

#### 4. DISCUSSÃO DOS RESULTADOS

Nesta parte do trabalho encontram-se as discussões sobre resultados das técnicas investigadas neste artigo a partir da aplicação do modelo aos dados da base. Foram aplicados para este estudo de caso informações de uma amostra de aproximadamente 300 mil clientes, utilizando dados de relacionamento de 48 meses de toda essa base para estimação dos parâmetros do modelo, sendo que, para a verificação do resultado foi adotado o primeiro semestre de 2016.

Através da Figura 2, é possível verificar a aderência do modelo referente à previsão na quantidade de clientes em relação à quantidade de meses ativos, ou seja, quantidade de meses que ele realizou pelo menos uma compra. Considere o mês 01 (um) quando o cliente realiza sua primeira recompra na loja, ou seja, não é um tempo contínuo considerado, mas sim a quantidade de meses que o cliente retornou à loja completando de tal modo ao menos uma compra.

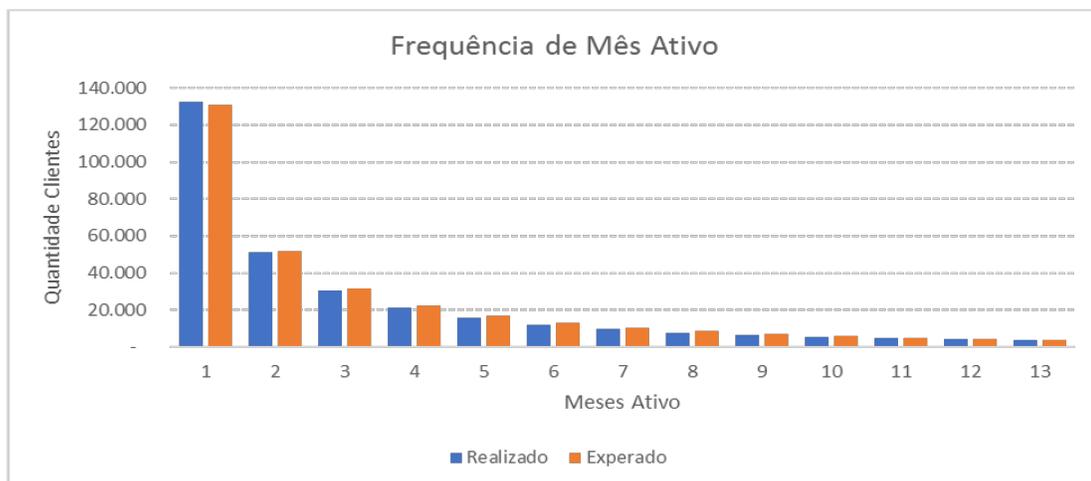


Figura 2: Frequência de transações

Fonte: O autor 2017.

Tabela 1: Frequência de transações em valores absolutos

<b>Meses Ativo</b>	<b>Realizado</b>	<b>Esperado</b>
<b>1</b>	132.399	131.132
<b>2</b>	51.429	51.888
<b>3</b>	30.368	31.476
<b>4</b>	20.942	22.110
<b>5</b>	15.720	16.608
<b>6</b>	12.111	12.977
<b>7</b>	9.683	10.399
<b>8</b>	7.611	8.483
<b>9</b>	6.664	7.002
<b>10</b>	5.568	5.828
<b>11</b>	4.781	4.900
<b>12</b>	4.226	4.151

Fonte: O autor 2017.

Desta forma, ao se analisar o comportamento de compra de 300 mil clientes, é possível perceber que do total da amostra, pouco mais de 132 mil clientes (44,13%) estiveram mais de um mês ativo na loja, da mesma forma, verifica-se que 51.429 pessoas (17%) registraram uma compra em um terceiro mês. Agora, avaliando o resultado da estimativa do modelo como um todo, o valor do erro absoluto entre o valor esperado e o realizado foi em média 4,7%.

Como primeiro resultado analisou-se a probabilidade dos clientes estarem ativos ou não em certo período de tempo. Consideram-se ativos, neste estudo, os clientes para os quais for obtida uma probabilidade maior ou superior a 0,7 - caso contrário, foram considerados inativos.

Constata-se, como demonstrado na Figura 3, que o percentual de ativos referente a amostra no qual o modelo foi aplicado está em torno de 31,8%. Por exemplo, em fevereiro de 2016 o modelo indicou e identificou os 35% de clientes que são considerados ativos neste período, ou seja, esses 35% possuem maior probabilidade de comprarem no mês analisado.

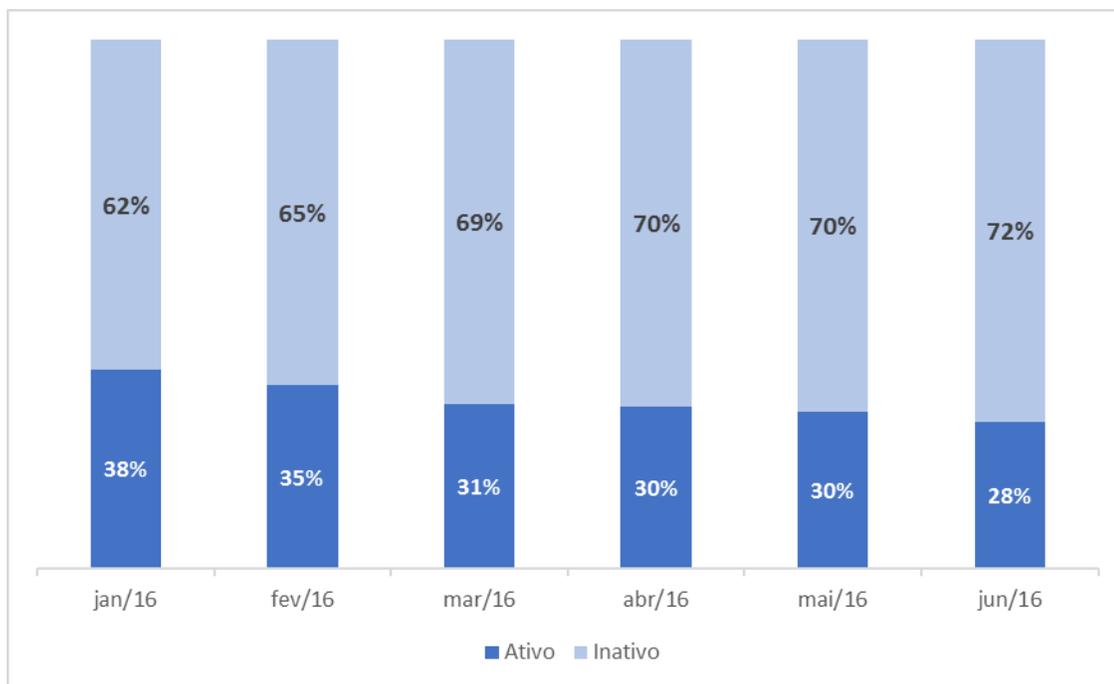


Figura 3: Clientes classificados com ativos/inativos ao longo dos meses de análise

Fonte: O autor 2017.

Porém, quando verificado do total de clientes que realmente realizaram um compra (reais ativos) no mês referente, dentro do grupo dos ativos, apurou-se um resultado médio de 86%, conforme demonstrado pela Figura 4.

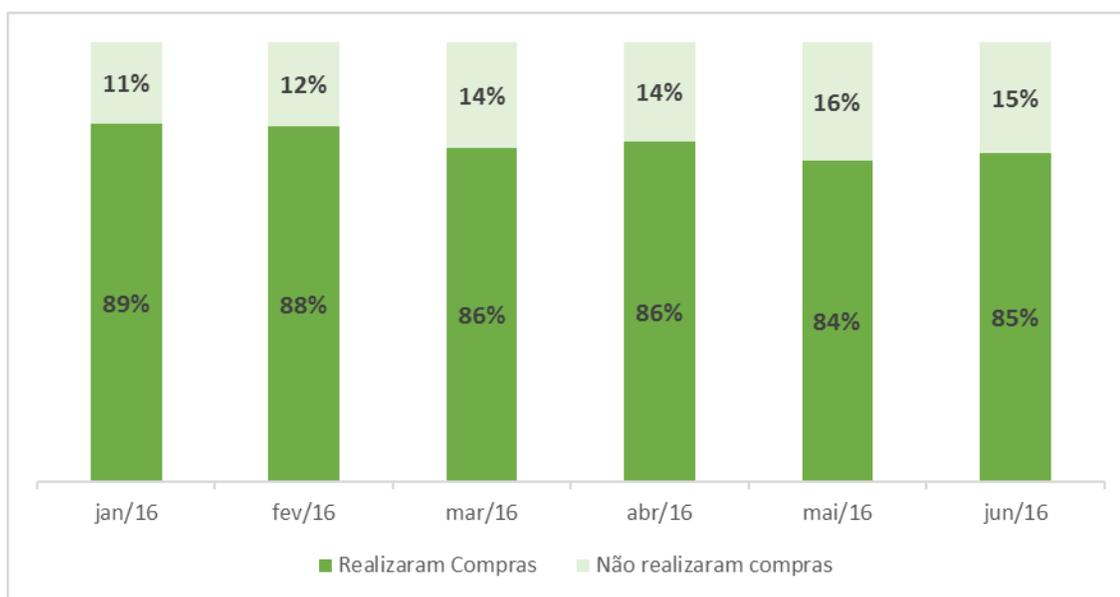


Figura 4: Cliente que realizaram ou não compras ao longo dos meses de análise.

Fonte: O autor 2017.

Ou seja, se analisarmos o mês de janeiro separadamente teremos o resultado conforme apresentado pela Figura 5

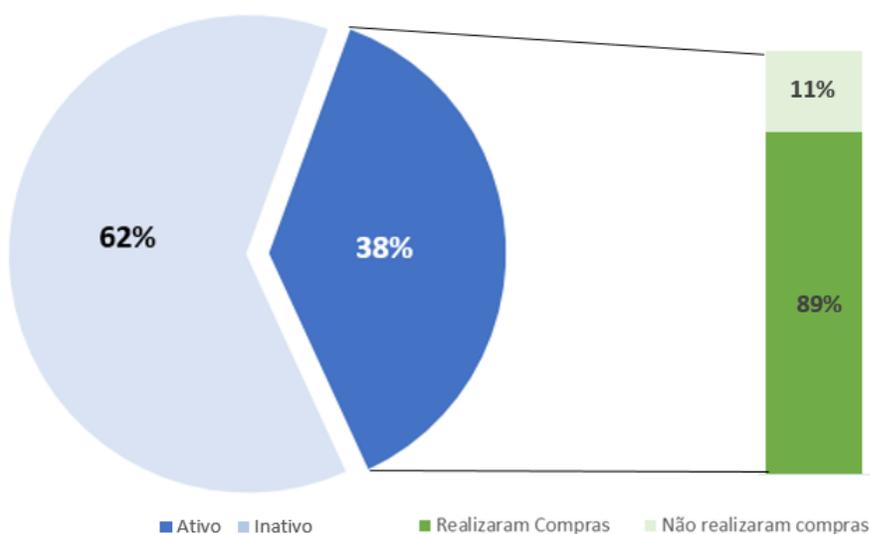


Figura 5: Cliente que realizaram ou não compras ao longo dos meses de análise, dentro do grupo de ativos – janeiro de 2016.

Fonte: O autor 2017.

O aspecto relevante derivado dos experimentos é o quanto de esforço pode ser direcionado com esses resultados para o trabalho marketing. Clientes podem abandonar uma organização ainda que altos investimentos em prospecção e retenção sejam realizados, o que requer diagnóstico e compreensão. Com esses resultados é possível direcionar tanto recursos e quanto esforços para tentar evitar a evasão desses consumidores.

Utilizando o resultado do DET, geraram-se as predições do tempo de vida dos clientes, e multiplicando este índice pelo lucro médio do cliente, é possível determinar uma das medidas mais importantes, o *Lifetime Value* do cliente.

Daí a análise da Figura 5, perceber-se que a distribuição de CLV gerada apresenta forte assimetria positiva. Verifica-se, por exemplo, que os 30% de maior contribuição respondem por mais de 87% do total do valor futuro gerado. Esse resultado nos revela a dimensão da importância dos maiores clientes para o faturamento da empresa.

Tabela 2: Classificação dos clientes em relação a faixa de CLV.

Faixa CLV		Percentil Clientes	CLV Acumulado	% LTV
De	Até			
Mínimo	R\$ 2	10%	R\$ 23.645	0%
R\$ 2	R\$ 6	10%	R\$ 74.795	0%
R\$ 6	R\$ 14	10%	R\$ 182.800	0%
R\$ 14	R\$ 29	10%	R\$ 405.661	1%
R\$ 29	R\$ 55	10%	R\$ 832.880	2%
R\$ 55	R\$ 98	10%	R\$ 1.512.846	4%
R\$ 98	R\$ 169	10%	R\$ 2.603.304	6%
R\$ 169	R\$ 295	10%	R\$ 4.495.983	11%
R\$ 295	R\$ 567	10%	R\$ 8.210.360	20%
R\$ 567	Máximo	10%	R\$ 23.576.081	56%

Fonte: O autor 2017.

Com base neste resultado é possível definir, de uma forma precisa, a receita que poderá ser gasta com cada grupo de cliente em ações de marketing, como retenção. É necessário alocar mais recursos em grupos de clientes que trarão um maior lucro para a organização. Já os grupos menos lucrativos devem ser trabalhados de uma maneira diferenciada para que de alguma forma possa transforma-los em clientes lucrativos e/ou fieis. Empresas que possuem este tipo de informação podem obter grande vantagem.

Os resultados indicam que os modelos de classificação por meio de mineração de dados desenvolvidos por esse tipo de análise, podem ser úteis na classificação dos grupos de clientes pela empresa e locação de recursos de maneira inteligente, dessa forma, é possível melhorar o desempenho da organização e até mesmo avaliar o valor de clientes individuais, determinando o seu tempo de vida como cliente e o lucro que proporcionará.

## **CONSIDERAÇÕES FINAIS**

O objetivo principal do trabalho é aplicar a ferramenta *Data Mining* para organização dos dados e gerar informação para tomada de decisão em marketing em uma empresa de *e-commerce* de vinho. Para o desenvolvimento foi utilizado um modelo probabilístico BG-BB, uma técnica não usual no *Data Mining*.

O processo de descoberta de conhecimento e o uso dos novos modelos baseados em Mineração de Dados desenvolvidos produzem reais práticas para uma loja de *e-commerce*. A utilização correta dessas variáveis ajudará em uma melhor classificação e previsão do comportamento dos seus clientes.

Sendo importante destacar que é preciso ter atenção caso o comportamento de compras dos clientes mude significativamente ao longo de tempo, os parâmetros do modelo devem ser revistos e logicamente a aderência do modelo. Além disso, a utilização de outros conjuntos de análises dará robustez à tomada de decisão das ações estratégicas.

Por fim, os resultados obtidos atenderam à expectativa de descoberta de padrões e comprovaram que essa técnica pode ser utilizada para auxílio na tomada de decisão desse tipo de empresa. Não basta dominar as técnicas para atrair novos clientes; a empresa deve retê-los. Muitas empresas possuem um alto índice de rotatividade de clientes, ou seja, conquistam novos clientes e perdem muitos deles. Muito pode ser feito com relação aos clientes utilizando esse tipo de metodologia, examinando constantemente percentuais de clientes que abandonam ou que estão comprando abaixo de níveis desejados para a corporação.

## REFERÊNCIAS

BATISTA, Pablo Estevan; FRAGA, Me Rinaldo Martins. Fidelização de clientes no mercado externo através de estratégias de relacionamento: estudo de caso de uma empresa de calçados da cidade de Três Coroas-RS. **Revista de Administração de Empresas Eletrônica-RAEE**, n. 3, 2016.

BRACHMAN, Ronald J.; ANAND, Tej. **The process of knowledge discovery in databases**. In: **Advances in knowledge discovery and data mining**. American Association for Artificial Intelligence, 1996. p. 37-57.

LINOFF, Gordon S.; BERRY, Michael JA. **Data mining techniques: for marketing, sales, and customer relationship management**. John Wiley & Sons, 2011.

CARDOSO, Olinda NP; MACHADO, Rosa TM. **Gestão do conhecimento usando data mining**: estudo de caso na Universidade Federal de Lavras. *Revista de administração pública*, v. 42, n. 3, p. 495-528, 2008.

DA COSTA, Luiz Carlos et al. **Data Mining....** Ou procurar uma pepita de ouro em uma poça de lama!. Disponível em: <[http://www.abepro.org.br/biblioteca/ENEGETP2000\\_E0229.PDF](http://www.abepro.org.br/biblioteca/ENEGETP2000_E0229.PDF)>. Data de acesso: (10 nov. 2016).

CORRÊA, Ângela Mc Jorge; SFERRA, H. H. Conceitos e aplicações de data mining. **Revista de ciência & tecnologia**, v. 11, p. 19-34, 2003.

ELMASRI, R.; NAVATHE, S. B. **Sistemas de Banco de Dados**. 4.ed. São Paulo: Pearson Addison Esley, 2005.

DA SILVA, Mabel Pereira; BOSCARIOLI, Clodis; PERES, Sarajane Marques. **Análise de logs da web por meio de técnicas de data mining**. 2003. Disponível em: <

[http://conged.deinfo.uepg.br/~iconged/Artigos/Artigo\\_03.pdf](http://conged.deinfo.uepg.br/~iconged/Artigos/Artigo_03.pdf)>. Data de acesso: (16 nov. 2016).

DE ARAÚJO RIBEIRO, Rodrigo Otávio; SANT'ANNA, Annibal Parracho. **Aplicação do modelo bg/bb na previsão do tempo residual de relacionamento de clientes.** Disponível em: <<http://www.din.uem.br/~ademir/sbpo/sbpo2007/pdf/arq0033.pdf>>. Data de acesso: (15 nov. 2016).

DE MORAES SOUSA, Marcos; FIGUEIREDO, Reginaldo Santana. **Análise de crédito por meio de mineração de dados:** aplicação em cooperativa de crédito. Revista de Gestão da Tecnologia e Sistemas de Informação, v. 11, n. 2, p. 379-396, 2014.

ENGEL, J. F.; BLACKWELL, R. D.; MINIARD, P. W. **Comportamento do consumidor.** 8. ed. Rio de Janeiro: Livros Técnicos e Científicos Editora S.A,2000.

FADER, Peter, HARDIE, Bruce e BERGER, Paul D. **Customer-base analysis with discrete-time transaction data.** 16.Tese - University of Auckland, 2004.

FADER, Peter S.; HARDIE, Bruce GS; LEE, Ka Lok. **Counting your customers the easy way:** An alternative to the Pareto/NBD model. Marketing science, v. 24, n. 2, p. 275-284, 2005.

FAYYAD, U. M. et al. **From data mining to knowledge discovery:** an overview. In: Advances in knowledge discovery and data mining. California: AAAI/The MIT, 1996. p.1-34.

FAYYAD, Usama; PIATETSKI-SHAPIRO, Gregory; SMYTHI, Padhraic. **The KDD Process for Extracting Useful Knowledge from Volumes of Data.** In: Communications of the ACML, pp.27-34, Nov.1996b.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa.** 5. ed. São Paulo: Atlas, 2008.

GUPTA, Sunil; NEHMANN, Donald R. **Gerenciando clientes como investimentos.** Bookman Editora, 2009.

KOTLER, Philip. **Administração de Marketing.** 4 ed. São Paulo: Atlas, 1996.

KOTLER, Philip. **Administração de Marketing.**10. ed. São Paulo: Person, 2000.

KOTLER, Philip. **Administração de Marketing:** análise, planejamento, implementação e controle. 5. ed. São Paulo: Atlas, 1998.

KOTLER, Philip; KELLER, Kevin Lane. **Administração de marketing.** 2006.

KUMAR, V. e REINARTZ, Werner e **Customer Relationship Management, A Database Approach**. John Wiley & Sons, June 2005.

MADRUGA, R. Guia de implementação de marketing de relacionamento e CRM. São Paulo: Atlas, 2006.

MARTINHAGO, Sergio. **Descoberta de Conhecimento sobre o Processo Seletivo da UFPR**. 125. Tese de Doutorado. Universidade Federal do Paraná, 2005.

MILAN, Gabriel Sperandio. **A prática do marketing de relacionamento e a retenção de clientes: um estudo aplicado em um ambiente de serviços**. 224. Tese de Doutorado - Universidade Federal do Rio Grande do Sul. Escola de Engenharia. Programa de Pós-Graduação em Engenharia de Produção. Porto Alegre, 2006.

QUEIROGA, Rodrigo Mendonça. **Uso de técnicas de data mining para detecção de fraudes em energia elétrica**. 147. Dissertação de Mestrado, Universidade Federal do Espírito Santo, Vitória-ES, 2005.

PEITER, Jorne Ricardo; DE MELO ABICHT, Alexandre. **A percepção do cliente em relação às ferramentas de marketing de relacionamento da Souza Cruz sa**. Dissertação de Mestrado, Universidade Cândido Mendes, Rio de Janeiro-RJ, 2005, p. 31.

RODRIGUES, Maria Carolina Bernardini. **O marketing de relacionamento como estratégia para a obtenção de vantagem competitiva**– um estudo de caso em uma pequena empresa. 2009. Disponível em: <<http://aberto.univem.edu.br/handle/11077/502>>. Data de acesso: (15 nov. 2016).

ROCHA, L. C. **Orientação para clientes**. Rio de Janeiro: Senac Nacional, 2004.

SANDRI, Samuel Diego. **Data mining–aplicação em uma base de dados real com dados de usuários dos faróis do saber** (doi: 10.5329/RESI. 2008.0702005). Revista Eletrônica de Sistemas de Informação ISSN 1677-3071 doi: 10.21529/RESI, v. 7, n. 2, 2009.

SWIFT, Ronald S. Accelerating customer relationships: Using CRM and relationship technologies. Prentice Hall Professional, 2001.

YIN, Robert K. **Estudo de Caso: Planejamento e Métodos**. São Paulo: 3 ed. Ed. Bookman, 2005.