

5

Sistema de Auxílio à Retenção de Clientes por Mineração de Dados

5.1.

Introdução

O *churn* em telefonia celular é um problema sério e com sérias conseqüências e custos, como já foi discutido. Fica claro que, dada a natureza da questão e os impactos nocivos da perda de clientes para um concorrente, a forma mais econômica de se enfrentar o problema é encontrar uma maneira de evitar que o abandono de clientes valiosos ocorra, ao invés de correr atrás de clientes perdidos e arcar com os grandes custos de aquisição e reaquisição de clientes, além de grandes gastos com imagem e publicidade [MATT01] [YAN04]. Para tanto, é necessário que uma operadora compreenda seus consumidores, conheça seus desejos e críticas, entenda seus comportamentos e, por fim, seja capaz de identificar quais dos seus milhões de clientes estão prestes a cometer *churn*. Somente sabendo, com alguma antecedência, quem são esses consumidores é que uma operadora pode agir de forma pró-ativa através de incentivos e promoções para evitar a perda do cliente. O conhecimento para alcançar esse objetivo reside nas inúmeras bases de dados operacionais e de relacionamento das empresas. Voltando sua atenção para elas, as operadoras podem adquirir a compreensão do cliente necessária para realmente prevenir o *churn*.

O sistema de retenção de clientes por mineração de dados para telefonia celular aqui apresentado visa exatamente a essa meta: extrair o conhecimento existente em bancos de dados de modo a tornar possível o entendimento do comportamento dos clientes e, conseqüentemente, a identificação prévia de quais deles seriam possíveis *churners* [ARCH04] [AU03] [BERR00] [MOZE00] [YAN04], gerando inteligência para ações de retenção bem direcionadas.

Apesar do sistema apresentado aqui ser voltado especificamente para a questão do *churn* em telefonia celular, acredita-se que ele apresente conceitos e elementos suficientemente amplos para ser aplicado em outras indústrias que enfrentem o *churn* (como bancos e seguradoras).

5.2.

O sistema de auxílio à retenção de clientes por mineração de dados

Um sistema realmente completo, que possa servir como base para ações e estratégias de retenção de clientes para uma empresa atuante no mercado de telefonia celular, é composto por diversas etapas.

Como apresentado na Figura 5.1, o ponto de partida são os dados operacionais e de relacionamento espalhados pela empresa, onde reside a informação. Em um cenário ideal estes dados já estariam organizados em *data warehouses* e *data marts*, o que facilitaria em muito sua obtenção, mas na prática tal organização ainda é rara de ser encontrada. Sob a ótica do problema do *churn* e do negócio de telefonia celular e suas peculiaridades, o passo inicial é definir que tipo de informação, dentre todos os dados disponíveis, poderia ser importante na caracterização do cliente objetivando sua retenção. Uma vez definidos claramente quais dados poderiam contribuir nesse sentido, é realizado um esforço para sua obtenção e agregação. Se algum dos dados desejados não estiver disponível, devem ser revisadas as necessidades iniciais, de forma a substituir a informação que não pode ser obtida por outra de natureza similar, ou até mesmo redefinir todas as necessidades de dados em virtude das limitações encontradas para adquiri-los.

Realizada a captura e consolidação dos dados a serem utilizados, segue-se com uma análise exploratória dos dados, visando ao entendimento da sua estrutura, ao mesmo tempo em que se busca sua validação através da procura por inconsistências e erros. Neste momento deve ser realizada a limpeza dos dados conforme descrito anteriormente, para garantir sua homogeneidade e confiabilidade. Se uma parte dos dados estiver incorreta e for impossível repará-la, deve ser realizada uma nova tentativa para sua obtenção.

Acreditando-se que os dados estejam livres de erros (pelo menos em sua grande parte), o próximo passo é a definição do alvo da análise a partir das regras do negócio e da empresa em particular. O alvo é a variável chave que dirá se o cliente presente na base de dados pertence ao grupo dos clientes ativos ou dos clientes que cometeram *churn*, permitindo então a futura compreensão do perfil de cada uma das classes.

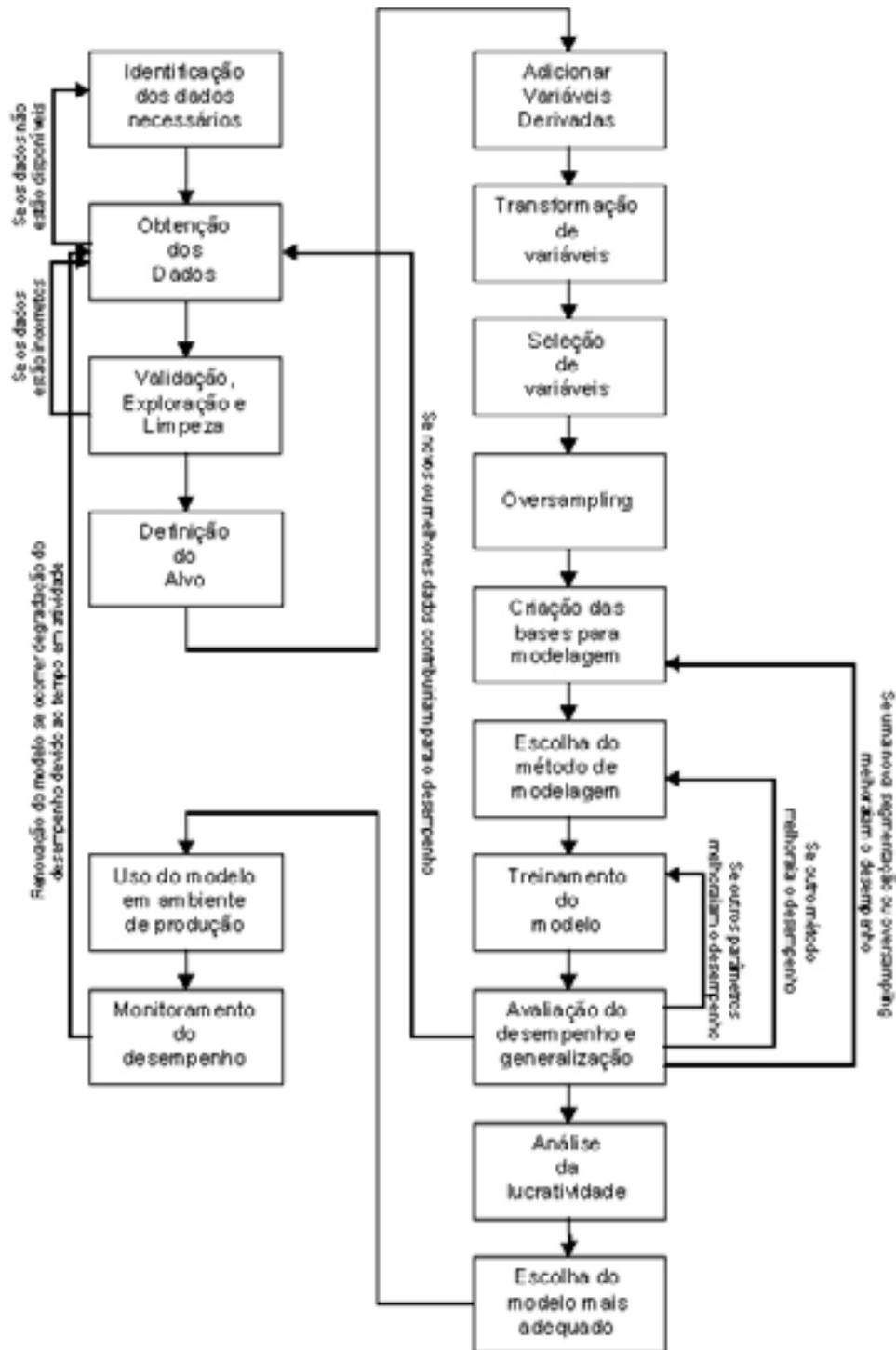


Figura 5.1 – Sistema de Auxílio à Retenção de Clientes por Mineração de Dados

As etapas seguintes à definição do alvo objetivam a criação das bases de dados para a posterior modelagem, incrementando o banco de dados através de: adição de variáveis secundárias, derivadas das existentes; transformações de variáveis [PYLE99]; e seleção das variáveis [BACK01] [BLUM97] [JANG93]

[KWAK03] [YI97] [YAN01] [ZHEN04] que melhor podem contribuir para a caracterização do alvo. Aqui também é realizado o *oversampling* da base [BERR00], de forma a obter bases cuja taxa de clientes *churners* (classe rara) seja suficientemente grande para ser reconhecida pelos modelos a serem aplicados.

Uma vez que as bases de dados estejam prontas e divididas em conjuntos de treinamento, validação e teste (para o estudo da generalização dos modelos), seleciona-se o modelo de classificação de padrões a ser usado para a tarefa de análise do perfil dos clientes ativos e *churners*, de forma a identificá-los corretamente. Tal modelo é então treinado buscando-se obter o melhor desempenho no problema de classificação de clientes. Vários tipos de modelos diferentes devem ser testados, objetivando encontrar aquele que melhor representa a realidade dos dados da empresa. Nota-se que uma análise da lucratividade de cada modelo para o mercado de telefonia móvel é realizada antes da definição de qual modelo é o ideal.

Finalmente, com o modelo mais adequado escolhido e tendo sua efetividade e lucratividade comprovadas, este deve ser aplicado ao ambiente de produção da empresa para realmente servir de fonte de informação sobre o cliente. Neste momento o sistema de auxílio à retenção de clientes por mineração de dados está inserido nas ações de retenção de clientes da empresa, sendo o mecanismo fundamental para a seleção de clientes que deverão receber incentivos ou observados para evitar o *churn*. Uma vez aplicado, o desempenho do sistema deve ser sempre monitorado, de forma que mudanças na empresa, ou na própria estrutura de mercado, não afetem seu desempenho. Se ocorrer degradação da resposta do sistema, principalmente devido à temporalidade dos dados e a alterações no que define o perfil e razões de um *churner*, uma nova obtenção de dados deve ser realizada e o sistema como um todo deve ser atualizado para corresponder à nova realidade.

A Figura 5.1 ilustra todas as partes e processos do sistema desenvolvido de retenção de clientes por mineração de dados. As seções seguintes apresentam, em mais detalhes, cada uma destas etapas.

5.3. Coleta e estudo dos dados

O primeiro passo para uma correta modelagem do problema do *churn* é uma coleta de dados adequada. Para coletar e agregar estes dados é

necessário primeiro entender a natureza dos dados geralmente disponíveis em uma operadora e definir quais deles serão necessários para o sistema [BERR00]. Em geral os tipos de dados presentes nas empresas de telefonia celular, e que são interessantes para o estudo do *churn*, podem ser agrupados da seguinte forma [MOZE00]:

- **Uso da rede:** detalhes sobre as chamadas feitas pelo usuário (data, tempo de duração, localidade, números de telefones diferentes discados, tipo de ligação), detalhes sobre chamadas perdidas (devido à falta de cobertura, por exemplo) e dados sobre a qualidade do serviço prestado (interferência, má cobertura);
- **Faturamento:** toda a informação financeira que aparece na conta do usuário (faturas mensais, valor da assinatura, cobranças por uso roaming, cobranças por minutos adicionais, entre outros);
- **Atendimento ao Cliente:** dados sobre os contatos feitos entre cliente e o serviço de atendimento ao cliente e suas resoluções;
- **Relacionamento:** detalhes sobre o relacionamento com o cliente (antiguidade da conta, aparelho possuído, tecnologia utilizada, situação de crédito, e outros);
- **Demográfico:** dados sobre a posição geográfica dos clientes e suas características como população (sexo, idade, estado civil, etc);
- **Mercado:** informação sobre o mercado e os competidores (custos de publicidade, tarifas da competição, etc);

Apesar de sempre se desejar ter acesso à maior quantidade de informação possível, muitas vezes alguns dos dados desejados não são confiáveis ou não podem ser obtidos. Deve-se trabalhar com as possibilidades permitidas pela realidade, mesmo sabendo que o sistema poderia ter um desempenho superior caso certos dados existissem. Nesse momento, o sistema de auxílio à retenção de clientes pode fornecer sugestões para a empresa sobre que tipos de dados não acessíveis deveriam estar disponíveis e de que forma, para que no futuro a realidade seja outra.

Outro detalhe importante na indústria de telefonia celular é a necessidade de segmentação dos dados para a elaboração dos modelos. Isso ocorre porque existem vários tipos de clientes e cada um deles possui razões que podem ser bem distintas na causa do *churn*. Além disso, cada tipo de cliente possui perfis muito diferentes, o que tornaria impossível qualquer tentativa de modelá-los em conjunto [BERR00]. Os tipos de clientes mais comuns são os pré-pagos, os pós-pagos e os clientes empresa. Podem e devem existir outros tipos, dependendo

especificamente de cada operadora e suas definições de negócio. Na etapa de exploração, validação e limpeza dos dados esse fato deve ser observado e os clientes pertencentes a diferentes tipos devem ser separados para a construção de um sistema de retenção específico para cada categoria de consumidor.

Ao se explorar os dados, várias análises de distribuição e consistência devem ser realizadas, de modo que se entenda o perfil do cliente e se corrijam erros que porventura existam nos dados, como por exemplo valores impossíveis e outliers ou campos com muitos valores ausentes.

Uma vez obtidos, estudados, aprovados e compreendidos, os dados devem ser preparados para a modelagem do *churn*.

5.4. Preparação dos dados

É essencial a busca por uma representação dos dados mais adequada, que maximize a informação presente na base de dados para a tarefa de identificação do cliente como *churner* [MOZE00] [PYLE99]. As seções a seguir descrevem os passos necessários, tomados nesta dissertação, para se alcançar esse objetivo.

5.4.1. Definição do alvo

Não existe inicialmente na base de dados uma variável que diga se um registro corresponde a um cliente *churner* ou não. A variável alvo essencial, que dá ao sistema a capacidade de diferenciar um cliente ativo de um que abandonou a empresa, deve ser criada baseada nas definições e regras de negócio de cada operadora [YAN04]. Cada companhia pode possuir visões bem diferentes sobre o que caracteriza a saída de um cliente da empresa, e essas visões devem sempre ser respeitadas. Por exemplo, uma operadora pode requerer somente uma ligação do cliente para terminar os serviços por completo, enquanto outra necessita de uma carta do assinante para considerar os serviços terminados.

Em geral o procedimento utilizado na definição da variável alvo correta se inicia com a escolha de uma janela de previsão de *churn*. A janela de previsão ideal pode variar de uma semana até vários meses, dependendo da conjuntura de mercado e da situação da operadora. Uma vez definido o tamanho da janela de previsão, é escolhido um momento do histórico de dados que servirá como

período de base. Tal período de base deve possuir o mesmo tamanho da janela de previsão (semana, mês, etc). É necessário garantir a existência de pelo menos um período de tempo de mesmo tamanho após o período base, pois será neste primeiro período que será verificado o abandono do cliente, segundo as regras de negócio da empresa. Todos os clientes que satisfizerem tais condições de término de contrato, definidas pela operadora neste período após o período base, serão considerados *churners*.

5.4.2. Adicionar variáveis derivadas

Neste momento pode ser interessante a criação de novas variáveis a partir das variáveis existentes na base de dados, de forma a enfatizar certos aspectos dos clientes que podem ser de grande ajuda na definição do *churn*. Existem muitas formas de se realizar esse incremento na informação presente na base de dados, que em última análise dependem intrinsecamente dos dados existentes e de características da empresa.

Alguns exemplos da criação de variáveis derivadas são [MOZE02] [YAN04]: o cálculo de razões entre variáveis, como por exemplo, o gasto médio por minuto de ligação; a criação de variáveis que demonstrem mais claramente variações de outras entradas ao longo do tempo, como crescimento da fatura ao longo de alguns meses; ou até mesmo a criação de variáveis que agregem diversas outras com pouca informação uma a uma, como por exemplo uma variável para descrever todas as outras relacionadas ao serviço de *Call Center*. Tal variável seria a soma de todas as ligações para o *Call Center*, não importando sua natureza. As ligações para *Call Center* normalmente estão dispostas por tipo de ligação (reclamação, informação e serviço) em variáveis distintas em bases de *churn*.

5.4.3. Transformação dos dados

O passo de transformação de dados é também de grande importância para um melhor desempenho dos modelos a serem executados. Assim como o processo de criação de variáveis derivadas, a transformação de dados depende diretamente dos dados disponíveis e é impossível ditar, de forma genérica, todas as transformações que poderiam vir a ser úteis no processo de melhoria da

representação dos dados sem estar com os dados da operadora específica em mãos.

De qualquer forma, os procedimentos de [PYLE99] e descritos brevemente em capítulos anteriores podem e devem ser aplicados aqui. Em, particular a normalização dos dados é sempre importante em problemas de *churn*, pois a variabilidade e valores de máximo e mínimo de cada variável podem ser muito distintos, ainda mais dadas as variadas origens dos dados, que podem vir desde sistemas de faturamento até sistemas de atendimento ao cliente. Outra transformação constantemente realizada é o mapeamento de variáveis categóricas com n classes em n variáveis *dummy* com um único valor não nulo indicando a classe à qual pertence o registro. Tal mapeamento será utilizado no capítulo 6 para tratar as variáveis de sexo e categoria de cliente, por exemplo.

5.4.4. Seleção de variáveis

O sistema de retenção de clientes desenvolvido aqui aplica os três métodos de seleção de variáveis detalhados no Capítulo 3 (o Método do Estimador de Mínimo Quadrado – LSE [CHUN00] [CONT02], o Método da Efetividade de uma Entrada Singular – SIE [CAO97] [CAO96-2] [CONT02] e o Método baseado no Modelo ANFIS [JANG93] [CONT02]) na tentativa de encontrar e reconhecer que variáveis presentes na base de dados contribuem mais significativamente para a definição do *churn*. Além de reduzir a dimensão do banco de dados através da eliminação de variáveis com pouca ou nenhuma relevância para caracterização do alvo, tal procedimento também fornece idéias claras sobre quais são os atributos que levam um cliente a abandonar a empresa. O produto desses resultados pode ser usado pela operadora na compreensão do cliente e no planejamento de incentivos mais eficazes.

5.4.5. Oversampling

Como descrito anteriormente, é essencial um procedimento de *oversampling* [BERR00] para que bases de dados de *churn* sejam tratáveis por modelos de classificação de padrões. Observando-se novamente a Tabela 4.2 com as taxas de churn mensais médias para as quatro principais operadoras do Brasil (2% para a Vivo, 3,1% para a Claro, 2,2% para a TIM e 1,8% para a Oi), nota-se com clareza que qualquer base de dados de qualquer operadora conterà

um número inferior a 4% dos registros pertencentes à classe de *churner*. Tal proporção deve ser corrigida através de *oversampling* para algum valor entre 10% e 40%, de modo que os modelos a serem aplicados consigam distinguir ambas as classes da variável alvo.

No sistema desenvolvido, devido à limitação do número de casos de *churners* presentes na base de dados disponível para testes, foi utilizada a proporção fixa de 28.6% de clientes da classe *churner* em todas as bases de treinamento, validação e teste. Segundo [ARCH04] [BERR00], taxas em torno desse nível apresentam os melhores resultados.

5.5. Modelagem do churn

Com as bases de dados prontas para a modelagem e divididas em conjuntos de treinamento (70%), validação (20%) e teste (10%), realiza-se então a avaliação de diversos modelos de classificação de padrões, na busca pelo modelo que realize a classificação mais adequada dos clientes e possua a melhor capacidade de generalização.

Os modelos utilizados e testados pelo sistema foram:

- Classificadores Bayesianos Naïve [DUDA00];
- Redes Neurais MLP e PNN [BISH96] [MOZE00] [ZHEN04];
- Árvores de Decisão [QUIN87] [BERR00];
- Algoritmos Genéticos, com o Rule-Evolver [LOP99-1] [LOP99-2];
- Sistema Neuro-Fuzzy Hierárquico BSP (NFHB) [GONÇ01] [SOUZ99];
- Máquinas de Vetor de Suporte [ARCH04] [ZHEN04] [DUDA00].

Nesta etapa devem ser realizados vários testes com os diversos parâmetros que cada família de modelos possui, de modo a se alcançar o modelo mais próximo do ótimo possível para os dados presentes. Por exemplo: em redes neurais varia-se os algoritmos de treinamento, a taxa de treinamento, a taxa de momentum, o *weight decay*, o número de neurônios na camada escondida; em máquinas de vetor de suporte varia-se o parâmetro de regularização *C*, o tipo de *kernel* e os parâmetros do *kernel*; no Rule-Evolver varia-se as taxas e tipos de mutação e crossover, as funções de avaliação, o tipo de recompensa entre outros; na Árvore de Decisão varia-se o algoritmo de treinamento, a profundidade máxima, o mínimo de padrões por folha, o mínimo de padrões necessários para uma partição ocorrer; no NFHB varia-se a taxa de decomposição que limita o crescimento da sua estrutura.

O desempenho dos modelos é avaliado principalmente segundo a matriz de confusão por eles fornecida, onde estão claros os erros de classificação para cada classe, além da análise de lucratividade que será detalhada na próxima seção.

Em todos os processos de modelagem é feito um procedimento de validação cruzada *10-fold*, para ajudar na eliminação de problemas relacionados a divisões tendenciosas entre os conjuntos de treinamento, validação e teste. A proporção entre *churners* e *ativos* em todas as bases é sempre mantida.

Alguns dos modelos utilizados geram regras do tipo SE-ENTÃO, entre eles as árvores de decisão, o Rule-Evolver e o NFHB. Tais regras oferecem interpretabilidade aos resultados, o que pode ser de muito valor na compreensão dos clientes e dos fatores que levam ao *churn*. Mesmo que o modelo com o melhor desempenho ao final do processo de modelagem não seja um destes, se o desempenho for próximo do ótimo é interessante avaliar as regras geradas por eles em busca de informação útil.

5.6. **Análise de lucratividade**

No sistema de auxílio à retenção proposto, a escolha do modelo ideal a ser implementado pela operadora na caracterização de seus clientes não é simplesmente resultado da análise da taxa de classificação correta que os modelos apresentam. Após terem sido otimizados, os melhores modelos de cada família de métodos são submetidos a uma análise da lucratividade que eles proporcionariam para a operadora caso fossem a base para suas ações de retenção.

O estudo desta lucratividade, ou corte de gastos, que a aplicação de um modelo de mineração de dados para identificação do *churn* resultaria é baseada em [MOZE00], onde é desenvolvida uma metodologia completa para a tradução do desempenho de modelos de classificação de padrões em possíveis lucros para uma operadora.

Conforme já foi dito anteriormente, ações e estratégias de retenção giram em torno do oferecimento de algum incentivo para garantir a permanência na empresa de algum cliente que porventura tenha sido identificado como provável abandono. Através de tal incentivo espera-se que haja a eliminação ou pelo menos uma grande redução das chances do cliente contatado realmente trocar a operadora por uma concorrente, evitando, portanto, o *churn*. A identificação de

tais clientes a partir da mineração de dados visa exatamente a minimizar os custos destas ações de incentivo, evitando o contato com os clientes errados que não responderiam aos incentivos, ou que nem possuíam a intenção de *churn*.

Obviamente, qualquer incentivo tem o seu custo ($C_{incentivo}$), mas esse custo é sempre muito mais baixo do que a perda de receita causada pelo *churn* de um cliente (C_{perda}). Em última análise, é exatamente essa grande diferença entre o $C_{incentivo}$ e C_{perda} que realmente valida qualquer tentativa de se combater o *churn*. Mesmo levando-se em consideração o importante fato de que nem todos os clientes a quem seja oferecido um incentivo irão desistir do *churn* (a redução na probabilidade de *churn* por assinante devido a um incentivo será chamada de $P_{incentivo}$), grandes economias podem ser alcançadas.

Para expressar o que acabou de ser dito em números, é necessário fixar uma janela de tempo na qual é esperado que um incentivo faça efeito sobre o cliente. Para o sistema em questão foi fixada uma janela de seis meses, a exemplo de [MOZE00], mas o conhecimento de cada operadora sobre a sua estratégia de incentivos é o que deve comandar essa escolha em uma aplicação real. Com essa janela de tempo fixa, é possível calcular C_{perda} como sendo a soma da fatura média do cliente ao longo da janela de tempo, acrescida de um custo fixo que representa o gasto necessário para se adquirir um novo cliente para o lugar do que acabou de abandonar a empresa.

Então, para calcular a economia de gastos resultante da aplicação de um modelo de classificação de padrões na identificação de *churners*, os valores acima, $C_{incentivo}$, C_{perda} e $P_{incentivo}$, são combinados com a matriz de confusão do desempenho de um modelo. Essa matriz de confusão detalha o acerto de classificação do modelo para cada uma das classes, assim como os seus erros, em geral em valores percentuais. No entanto, para o cálculo da lucratividade é necessária uma matriz de confusão que possua o número de clientes classificados pelo modelo, correta ou erradamente, em cada um dos campos. O número absoluto de clientes dependerá da operadora e do universo definido por ela para a aplicação do sistema de retenção. Os elementos desta matriz seguem a seguinte notação:

- $n(pC, rC)$ é o número de clientes preditos *churners* (pC) e realmente *churners* (rC).
- $n(pC, rA)$ é o número de clientes preditos *churners* (pC) mas que na realidade são ativos (rA) – erro tipo 1, falso positivo.

- $n(pA, rA)$ é o número de clientes preditos ativos (pA) e realmente ativos (rA).
- $n(pA, rC)$ é o número de clientes preditos ativos (pA) mas na realidade *churners* (rC) – erro tipo 2, falso negativo.

Dadas estas definições, é necessário computar duas estatísticas: o valor do custo resultante para a operadora se ela não tomar nenhuma atitude para evitar a perda dos clientes identificados como possíveis *churners* e simplesmente deixá-los ir (V_{perda}); e o custo de oferecer incentivos para cada um dos consumidores identificados pelo modelo como *churners* ($V_{incentivo}$). Esses cálculos são feitos conforme descrito nas Equações 5.1 e 5.2.

$$V_{perda} = [n(pC, rC) + n(pA, rC)] \times C_{perda} \quad \text{Equação 5.1}$$

A Equação 5.1 mostra que, não importando se o *churn* é predito corretamente ou não pelo modelo, o cliente irá deixar o negócio a um custo de C_{perda} por assinante, resultando em uma perda total de V_{perda} .

$$V_{incentivo} = [n(pC, rC) + n(pC, rA)] \times C_{incentivo} + [P_{incentivo} \times n(pC, rC) + n(pA, rC)] \times C_{perda} \quad \text{Equação 5.2}$$

A Equação 5.2 diz que o custo de se oferecer um incentivo $C_{incentivo}$ será gasto para cada assinante que foi predito como *churner* pelo modelo, mas ao mesmo tempo ocorrerá uma redução de fração $P_{incentivo}$ nas perdas causadas pelos clientes corretamente classificados com *churners*.

Com esses resultados, é possível, finalmente, calcular a economia de custos estimada por cliente baseado em um modelo de classificação de padrões, segundo a Equação 5.3.

$$economia = \frac{V_{perda} - V_{incentivo}}{[n(pC, rC) + n(pA, rC)]} \quad \text{Equação 5.3}$$

A equação 5.3 deixa claro que a economia resultante do processo é exatamente a diferença entre o que a operadora gastaria não intervindo de forma alguma, deixando todos os *churners* abandonarem seus serviços, e o custo de se agir sobre os clientes previstos como *churners* pelo modelo. O denominador

da equação contendo o total de *churners* no universo de clientes avaliado serve somente para dar a estimativa de economia por cliente de provável abandono.

Com esse procedimento é possível então calcular a economia gerada por cada modelo que apresente resultados promissores, tendo uma visão de seu desempenho diretamente ligada ao problema do *churn* e ao seu impacto se aplicado como base para estratégias de retenção.

5.7. Operacionalização dos resultados

Depois de otimizado e selecionado o modelo ideal para a tarefa de entendimento do comportamento dos clientes e identificação dos mesmos como ativos ou possíveis *churners*, o sistema deve ser levado para um ambiente de produção na empresa. Lá ele desempenhará seu papel de classificação na frequência desejada pela empresa e gerará inteligência sobre os clientes na qual a operadora deve pautar suas ações de marketing, publicidade e atendimento ao consumidor com o objetivo de prevenir o abandono daqueles clientes considerados valiosos e evitar os gastos subseqüentes do *churn*.

Dado o dinamismo do mercado de telefonia celular, é necessário o monitoramento do desempenho do sistema e, preferencialmente, a sua atualização constante com o surgimento de novos dados. Desta forma, garante-se que as razões vigentes para o *churn* no momento de sua utilização fazem parte de sua operação e que o sistema não está preso a perfis de *churners* que com o passar do tempo não existem mais.

5.8. Resumo

Neste capítulo foi detalhado o sistema desenvolvido de retenção de clientes por mineração de dados, apresentando todas as suas etapas.

O próximo capítulo apresenta um estudo de caso da aplicação do sistema sobre uma base de dados real e sua eficácia.