

IDENTIFICAÇÃO DE NICHOS E DEFINIÇÃO DE PERFIS DE COMPORTAMENTO ATRAVÉS DE TÉCNICA DE CLUSTERING COM MAPA DE KOHONEN

Galo Carlos Lopez Noriega

Ibmec São Paulo

R. Quatá 300 – VI. Olímpia CEP: 04546-042

São Paulo-SP

<http://www.ibmecsp.edu.br>

Helena Gottschalk

Ibmec São Paulo

R. Quatá 300 – VI. Olímpia CEP: 04546-042

São Paulo-SP

<http://www.ibmecsp.edu.br>

Resumo

Hoje no planejamento estratégico de Marketing faz-se necessário entender o perfil do consumidor para que se defina, público alvo, produtos e canais de distribuição que gerem maior rentabilidade para a Empresa e maior satisfação para o consumidor. Dependendo do número e diversidade de clientes de uma empresa, simplesmente explorar a base de clientes através de análise descritiva não é suficiente. Partindo-se do princípio que o mercado de massas não explica o comportamento individual do cliente, uma boa aproximação do comportamento dele é a identificação de seu comportamento através de técnicas de clustering e depois traçar estratégias de marketing segmentadas.

Neste contexto um banco de Varejo decidiu desenvolver um projeto de segmentação comportamental que identifica grupos de clientes com perfis de comportamento específicos na forma do uso dos canais e produtos/serviços. A técnica utilizada foi a de Redes Neurais com mapas de Kohonen com o software SAS.

Palavras-chave: redes neurais; análise de *clustering*; indústria financeira

Abstract

Nowadays in the Marketing arena is fundamental to understand the profile of the consumer you are dealing with. Only after that is possible to define the consumer target strategy, by understanding his most valuable products and services needs. It is also important to understand how the customer interacts with the company. That strategy generates the greater Company's profit and greater customer's satisfaction. Depending on the number and the diversity of customers of a company, simply to use descriptive analysis is not enough to understand the customer specific needs. Assuming that the Mass market does not explain the individual customer behavior, a good approach is to identify the customer behavior through clustering techniques and later on to develop segmented strategies. In this context a retail bank decided to develop a project of marketing segmentation to identify specific customer profiles. Neural networks with Kohonen maps were the clustering technique chosen.

Keywords: neural networks; clustering analysis; financial industry

1. INTRODUÇÃO

Hoje no planejamento estratégico de Marketing faz-se necessário entender o perfil do consumidor para que se defina, público alvo, produtos e canais de distribuição que gerem maior rentabilidade para a Empresa e maior satisfação para o consumidor. Dependendo do

número e diversidade de clientes de uma empresa, simplesmente explorar a base de clientes através de análise descritiva não é suficiente. Partindo-se do princípio que o mercado de massas não explica o comportamento individual do cliente, uma boa aproximação do comportamento dele é a identificação de seu comportamento através de técnicas de segmentação (clustering) e depois traçar estratégias de marketing segmentadas.

Neste contexto um banco de Varejo decidiu desenvolver um projeto de segmentação comportamental que identifica grupos de clientes com perfis de comportamento específicos na forma do uso dos canais e produtos/serviços. A técnica utilizada foi a de Redes Neurais com mapas de Kohonen com o software SAS Enterprise Miner. Este software roda em um servidor UNIX acessando um banco de dados relacional em DB2 com todas as informações de cadastro e movimentações do Banco.

2. INFORMAÇÕES COLETADAS

Para o entendimento do perfil do consumidor é importante coletar informações de uso e movimentações nos principais canais de distribuição e de movimentações de produtos e serviços bancários. Segue a relação das variáveis de canais de distribuição e produtos mapeados (todas com base mensal):

- **Boca do caixa e Caixa expresso (Agência):** quantidade, de valores e indicador de uso do canal;
- **ATM e Telefone (Call Center e Telemarketing):** quantidade de vezes que o cliente utilizou o canal no mês.
- **Internet Banking:** quantidade de acessos, total de consultas e volume transacionado no Canal;
- **Cartão de Credito:** valor da fatura, forma de pagamento, pagamento de juros;
- **Cartão de Debito:** valor mensal gasto por estabelecimento;
- **CDC, Leasing, Consorcio de Automóvel, Empréstimo Pessoal Parcelado:** valor total do empréstimo, saldo devedor e numero de parcelas contratadas e pagas;
- **Cheque Especial:** índice de utilização e limite
- **Debito automático:** quantidade de débitos automáticos por categoria;
- **Fundos, Renda Fixa e Poupança:** quantidade aplicada inicialmente, resgatada e aplicada mensalmente;
- **Capitalização:** quantidades, valores aplicados e numero de parcelas;
- **Seguros:** quantidades por tipo de seguro.
- **Tarifas:** pacote/cesta de tarifa e tarifas avulsas.

Estas informações são coletadas e armazenadas mensalmente numa estrutura de Datawarehouse (DB2/UNIX). Um filtro que adicionamos foi de que o cliente deveria estar a pelo menos 6 meses no Banco e sua conta com movimentação mínima fixada pelo Banco como conta ativa. O histórico utilizado foi de 12 meses para clientes Pessoa Física, com mais de 6 meses como correntista.

3. TECNICA UTILIZADA

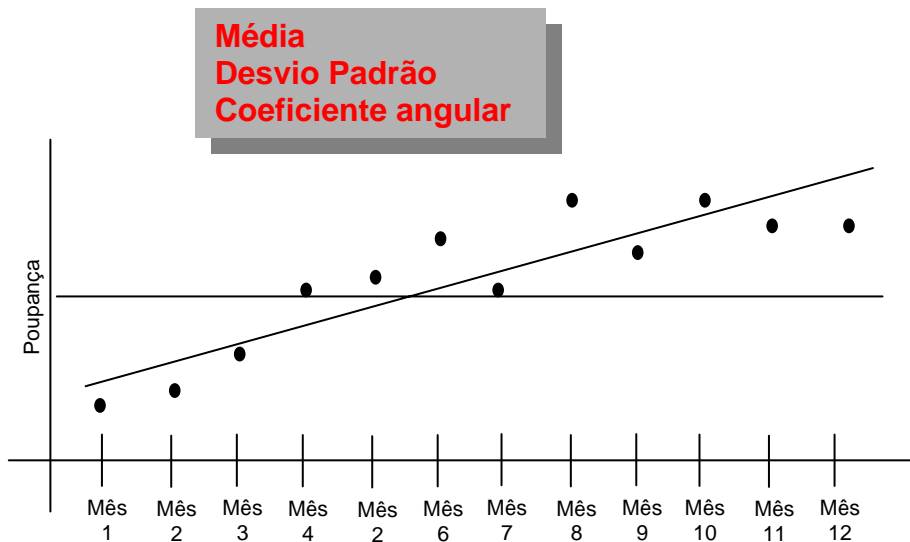
3.1 Análise de Clustering

As técnicas de análise de Clustering são utilizadas para a identificação de grupos de afinidade específicos numa determinada população, a função objetivo minimiza similaridade e maximiza diferenças (distancia euclidiana) para identificar os grupos através das variáveis de decisão previamente estabelecidas. Neste caso as variáveis de decisão são as de uso e movimentações de canais e produtos do banco. Existe uma serie de algoritmos de Clustering na literatura, mas resolvemos testar os que o software da empresa nos disponibilizava, K-Means e variações dos Mapas de Kohonen (algoritmos baseado em redes neurais).

3.2 Criação do vetor de comportamento do Cliente

A partir das variáveis coletadas no Datawarehouse com as informações de utilização e movimentação dos clientes do banco construímos um vetor de decisão que será dado como entrada ao algoritmo de clustering. Todas as variáveis dos clientes devem ser tratadas de forma a se construir um vetor de comportamento que represente uma síntese do perfil do cliente do banco, isto é, poderíamos, por exemplo, representar um cliente pela quantidade de vezes que ele vai a agência por mês, pelo total investido em fundos e pelo total de cheques que ele recebe. Começamos nosso estudo com quase 100 variáveis de utilização de canais e produtos, cada variável com um histórico de 12 meses (variável total investido por mês * histórico de 12 meses); então no total temos 100*12 variáveis a serem tratadas. O primeiro tratamento é uma compactação temporal das variáveis, do histórico de 12 meses de uma variável continua aplicamos uma transformação que calcula media, desvio padrão e tendência (coeficiente angular de uma regressão linear):

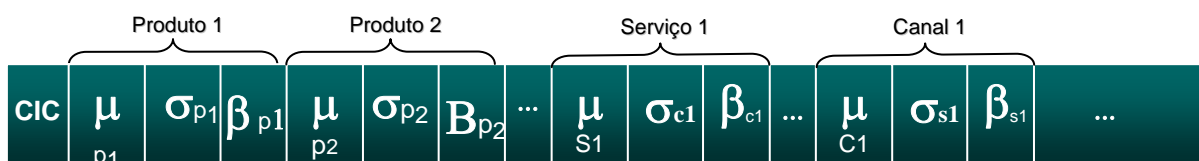
Exemplo: Produto Poupança



Neste exemplo o valor total mensal aplicado em poupança é transformado em Media, Desvio e Tendência do valor aplicado em poupança no ano.

Para todas as variáveis continuas (valores totais aplicados, resgatados, valor da fatura, saldo devedor, quantidades de acesso) esta transformação é aplicada. Para as variáveis discretas (indicador de uso de canais) foi utilizada a media do ano.

Segue abaixo uma ilustração do vetor de comportamento por cliente (CIC – código identificador do cliente) criado:



Onde: μ_i = Média do produto(p)/serviço(s)/canal(c) i σ_i = Desvio Padrão do produto/serviço/canal i β_i = Coeficiente Angular do produto/serviço/canal i

Deste vetor inicial foi feita uma análise de correlação e de significância para a redução de dimensão, chegando a 53 variáveis.

3.3 Metodologia - SOM (Self-Organizing Maps) / Kohonen

SOMs de Kohonen é um tipo de treinamento não supervisionado que tem como objetivo descobrir alguma estrutura subjacente (não conhecida) dos dados. Um mapa topológico é simplesmente um mapeamento que preserva as relações de uma vizinhança, minimizando as distancias para vetores de dados similares e maximizando a distancia para vetores distintos.

Segue abaixo as parametrizações feitas:

Variáveis de Entrada: 53

Tamanho Amostral para Segmentação: 100.000

Método de padronização de Variáveis: Amplitude $[(X_i - \text{Min})/(\text{Max} - \text{Min})]$

Regra de Agrupamento : ID - neste caso o CIC (código de Identificação do Cliente)

Dimensão do Mapa : 3 linhas X 4 colunas - Total de 12 clusters

Taxa de Aprendizado (α) : inicial = 0.9 final = 0.2

Passos até atingir a taxa final: 2.000 (I)

Max. Interações: 5.000 (II)

Max. Passos: 10.000.000 (I x II)

Critério de Convergência: 0.001

Vizinhança:

tamanho inicial: 5

tamanho final: 0

passos para atingir o tamanho final: 10.000

interações para atingir o tamanho final: 10

Sementes:

Método de Seleção Inicial: CP

Distância Min. entre Sementes: 0

Outras dimensões do Mapa de Kohonen foram testadas, (3x3 - 9 clusters) e (4x4 - 16 clusters). O critério de escolha do modelo final foi uma combinação de número de nichos (clusters) possíveis de serem tratados como estratégicos para a área de Marketing, diferenças relevantes encontradas em cada nicho, dispersão dentro do nicho sempre menor que a da população.

4. IDENTIFICAÇÃO E BREVE DESCRIÇÃO DOS NICHOS

A otimização não supervisionada resultou em 12 nichos com perfis de comportamento descritos abaixo:

Cluster/Nicho 1 – Primeiro Banco

Clientes muito bem relacionados com o banco

- Tempo de conta em torno de 10 anos
- Índice de Utilização do cheque especial em torno de 25%
- Limite Médio do Cheque Especial = 7400
- Limite Médio Cartão Crédito = 8300
- Rotativo em torno de 25%
- Saldo aplicado em Fundos e Poupança – 4 vezes maior do que a População
- Empréstimo CDC ou Leasing – 3 vezes maior que a População
- Capitalização – valor aplicado 10 vezes maior que a População
- Usa bastante agência, com tendência de migração para Internet Banking
- Neste nicho os outros produtos e serviços são utilizados conforme a População.

Cluster/Nicho 2 – PJ disfarçado (Comercio)

Clientes que movimentam sua conta como conta pessoa jurídica e tem atividade de pequeno comercio, por exemplo lojista, recebem e passam muitos cheques, transacionam bastante na agencia e por telefone, e usam produtos de credito.

- Tempo de conta maior que 7 anos
- Índice de Utilização do cheque especial maior que 37%
- Limite Médio do Cheque Especial = 4400
- Limite Médio Cartão Crédito = 4000
- Rotativo maior que 47%
- Usa empréstimo pessoal muito acima da media da população, tem varias parcelas;
- Nicho que mais tem produtos de Seguros, Debito Automático e Capitalização contratados
- Grande movimentação de cheques – 4 vezes mais que a População
- Utiliza muito agencia, Caixa Expresso e ATM
- Maior usuário do Call Center e mais acessado pelo Telemarketing

Cluster/Nicho 3- Moderno

Clientes que usam bastante o canal internet banking e faz uso racional dos produtos do banco.

- Tempo de conta menor que 3 anos
- Índice de Utilização do cheque especial abaixo de 37%
- Limite Médio do Cheque Especial = 3350
- Limite Médio Cartão Crédito = 3600
- Rotativo abaixo de 35%
- 2º Maior usuário do Internet Banking - 5 vezes mais que a População
- Boa Participação em passivos, previdência e cartão, com tendência de aumento.
- Uso na media da População dos produtos de empréstimos parcelados

Cluster/Nicho 4- Aposentado

Clientes que possuem produtos de empréstimos para aposentados e/ou que recebem seu beneficio do INSS no banco.

- Tempo de conta maior que 5 anos
- Índice de Utilização do cheque especial abaixo de 45%

- Limite Médio do Cheque Especial = 1000
- Limite Médio Cartão Credito = 1400
- Uso intensivo do Cartão de credito e de debito;
- Baixo volume aplicado em Fundos e Poupança;
- 2º Maior usuário de agencia e Caixa expresso.

Cluster/Nicho 5- PJ serviços altamente remoto

Clientes que movimentam sua conta como conta pessoa jurídica e tem atividade de serviços, por exemplo medico, recebem muitos cheques e transacionam bastante na internet.

- Tempo de conta abaixo de 4 anos
- Índice de Utilização do cheque especial em acima de 30%
- Limite Médio do Cheque Especial = 4000
- Limite Médio Cartão Credito = 4100
- Rotativo abaixo de 34%
- Usa muita tarifa avulsa;
- Boa Participação em Empréstimo Pessoal;
- Alto volume aplicado em Fundos e Poupança;
- Alto volume em Previdência com tendência de aumento;
- Muitos lançamentos na conta e muitos cheques compensados;
- Maior usuário do Internet Banking - 17 vezes mais que a população
- Alta movimentação em conta corrente.

Cluster/Nicho 6 – Renda complementar com produtos de Credito

Clientes que complementam mensalmente a renda usando os produtos de credito do banco de forma racional e saudável.

- Tempo de conta em torno de 10 anos
- Índice de Utilização do cheque especial maior que 44%
- Limite Médio do Cheque Especial = 1650
- Limite Médio Cartão Credito = 1700
- Rotativo em torno de 50%
- Usa muito cheque especial;
- Boa Participação em Empréstimo Pessoal;
- Alta movimentação no cartão de debito
- Utiliza muita agência, ATM e Telefone;
- Único com transação positiva no Boca de Caixa.

Cluster/Nicho 7 – Investidor Humilde

Clientes com poucos recursos, investem e deixam o dinheiro parado sem quase movimentar a conta corrente.

- Tempo de conta em torno de 6 anos;
- Índice de Utilização do cheque especial abaixo de 8%
- Limite Médio do Cheque Especial = 3200
- Limite Médio Cartão Credito = 2400
- Quase não usa os canais do Banco;
- Boa Participação em: Fundos (3.800), CDB (2.000), Poupança (2.200); com tendência de aumento;
- Baixíssima movimentação na conta corrente - 50% menos que a média do banco.

Cluster/Nicho 8 – Massa conveniência Cartão

Clientes com comportamento similar ao da população a menos do uso intensivo e crescente do cartão de crédito.

- Tempo de conta maior que 7 anos
- Índice de Utilização do cheque especial abaixo de 33%
- Limite Médio do Cheque Especial = 2500
- Limite Médio Cartão Crédito = 2400
- Rotativo acima de 30%
- Fatura cartão acima da média, com forte tendência de crescimento;
- Baixa utilização do Internet Banking;
- Neste nicho os outros produtos e serviços são utilizados conforme a População

Cluster/Nicho 9 – Folha de Pagamento não conquistada

Clientes com conta salário no banco mas que mantem vínculos com outros bancos.

- Tempo de conta maior que 5 anos
- Índice de Utilização do cheque especial abaixo de 5%
- Limite Médio do Cheque Especial = 2880
- Limite Médio Cartão Crédito = 2270
- Uso do Rotativo em torno de 27%
- Baixo uso do Cheque Especial – somente 25% usam;
- Produtos de Fundos e Poupança na média da População
- Utiliza somente cartão débito;
- Quantidade de Tarifas avulsas acima da media;
- Alta utilização de Caixa Automático;

Cluster/Nicho 10 – Abandonante

Clientes que já tiveram um bom relacionamento com o banco mas que ao longo do tempo vem diminuindo a tendência de valores e uso na maior parte dos produtos.

- Tempo de conta maior que 5 anos
- Índice de Utilização do cheque especial em media de 15%
- Limite Médio do Cheque Especial = 2260
- Limite Médio Cartão Crédito = 2240
- Rotativo menor que 10%
- Deixando de transacionar com o banco;
- Utilização de produtos e canais com tendência negativa.

Cluster/Nicho 11 – Pendurado

Clientes que estão bastante endividados mas ainda mantem um perfil saudável de crédito.

- Tempo de conta em torno de 3 anos;
- Índice de Utilização do cheque especial em media de 87% ;
- Limite Médio do Cheque Especial = 1160
- Limite Médio do Cartão Crédito = 1140
- Uso do Rotativo chegando a 72%
- Tendência de aumento de uso do cheque especial;

- Produtos de empréstimo pessoal bem acima da media da População.

Cluster/Nicho 12 – Devedores em renegociação

Clientes com grave problema de credito, alguns já com restritivos (divida sendo renegociada ou em cobrança)

- Tempo de conta em torno de 2 anos;
- Índice de Utilização do cheque especial > 90%
- Limite Médio do Cheque Especial = 700
- Limite Médio Cartão Credito = 870
- Uso do Rotativo chegando a 82%;
- Média de 29 restritivos.

Toda a base de clientes foi classificada nestes 12 nichos e acrescentada as informações demográficas foi possível concluir que alguns nichos de comportamento tem características demográficas fortes e especificas como idade, ocupação, faixa de renda, estado civil, região, etc. enquanto outros independem das características demográficas (como o Nicho 1 Primeiro Banco).

5. CONCLUSÕES

Este trabalho possibilitou a área de planejamento estratégico de marketing a entender melhor o relacionamento de seus clientes com o banco, inclusive quebrando paradigmas internos. Foram realizados vários projetos piloto, uns com foco em programas de cobrança, outros com programas de retenção, e a criação de pacotes de tarifas e cestas de produtos e serviços específicos para cada nicho ou grupos de nichos.

6. BIBLIOGRAFIA

- [1] KOHONEN, T. Self-Organizing Maps. 3. Auflage Springer, Berlin, 1995, ISBN 3-540-67921-9
- [2] OBERMAYER, K. Self-Organizing Map Formation: Foundations of Neural Computation
- [3] KOHONEN, T .Self-Organizing Maps. Springer-Verlag New York, Inc. 2001
- [4] PEPPERS, D.; ROGERS, M. Managing Customer Relationships: A Strategic Framework
- [5] GOTTSCHALK, H.; KIM, H.; ZWIETERING, D.; REINSCHIMDT, J. Intelligent Miner for Data: Enhance Your Business Intelligence, IBM Corporation, US, 1999
- [6] KOTLER, P.; ARMSTRONG, G. Princípios de Marketing. Pearson, Prentice Hall. 9ª Edição.
- [7] BISHOP, C. Neural Networks for Pattern Recognition Oxford University Press New York, Inc.1995
- [8] HUGHES, A. Strategic Database Marketing McGraw-Hill, 1994