

**UNIVERSIDADE FEDERAL DO MARANHÃO
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE ELETRICIDADE
ÁREA DE CONCENTRAÇÃO: CIÊNCIA DA COMPUTAÇÃO**

**MODELAGEM COMPORTAMENTAL DE USUÁRIOS: ABORDAGEM AOS
CONTRIBUINTES FAZENDÁRIOS**

**SÃO LUIS
2006**

MARCELO LUÍS LOBATO NOGUEIRA

**MODELAGEM COMPORTAMENTAL DE USUÁRIOS: ABORDAGEM AOS
CONTRIBUINTES FAZENDÁRIOS**

Dissertação apresentada como requisito parcial para obtenção de grau de Mestre em Engenharia da Eletricidade, com área de concentração em Ciência da Computação pelo Programa de Pós-Graduação em Engenharia de Eletricidade da Universidade Federal do Maranhão.

Orientador: Prof. Dr. Sofiane Labidi

SÃO LUIS
2006

MODELAGEM COMPORTAMENTAL DE USUÁRIOS: ABORDAGEM AOS CONTRIBUINTES FAZENDÁRIOS

Dissertação apresentada como requisito parcial para obtenção de grau de Mestre em Engenharia da Eletricidade, com área de concentração em Ciência da Computação pelo Programa de Pós-Graduação em Engenharia de Eletricidade da Universidade Federal do Maranhão.

Aprovada em 29/11/2006.

BANCA EXAMINADORA

Prof. Dr. Sofiane Labidi
Universidade Federal do Maranhão – UFMA
Orientador

Prof. Dr. Zair Abdelouahab
Universidade Federal do Maranhão – UFMA
Membro Interno

Prof. Dr. Renato da Veiga Guadagnin
Universidade Católica de Brasília – UCB
Membro Externo

Nogueira, Marcelo Luís Lobato.

Modelagem comportamental de usuários: abordagem aos contribuintes
fazendários / Marcelo Luís Lobato Nogueira. – 2006.

79 f.

Orientador: Sofiane Labidi

Dissertação (Mestrado) – Programa de Pós-Graduação em Engenharia de
Eletricidade, Universidade Federal do Maranhão, 2006.

1. Modelagem de sistemas (Computação). 2. Modelagem de usuários
(Computação). 3. Modelagem comportamental (Computação). I. Título.

CDU: 004.414.23

A meus pais, Joaquim e Flora, a
minha esposa, Tatiana, a minha
linda filha, Aline, e a minha tia
Vera Lúcia com carinho.

AGRADECIMENTOS

A Deus, por me conceder o dom da vida e me guiar no meu caminho.

A meus pais Joaquim e Flora, pelo amor e por sempre acreditarem em mim.

A meus irmãos, pelo carinho e ajuda ao longo da minha vida.

A minha tia Vera Lúcia, pela constante ajuda ao longo da minha formação acadêmica.

A minha esposa Tatiana, pela compreensão da minha dedicação aos estudos.

A minha filha Aline, pelos maravilhosos sorrisos nas horas difíceis.

Ao Prof. Dr. Sofiane Labidi, pela orientação e incentivo no decorrer deste trabalho.

A todos os amigos da SEFAZ-MA que contribuíram para a conclusão deste trabalho.

“Desvia-te do mal, e faze o bem; e
terás uma morada eterna”

Salmos 25:27

ABSTRACT

In this work, a generic user model, obtained from historical data from actions of the users and inference functions of belief is proposed. The functions of belief are obtained through forecasting techniques. This model was tested with its utilization in taxpayers. Two forecasting techniques had been used: statistical Holt model and neural networks. The results obtained from tests in Government of Maranhão, in Brazil, are also reported.

Keywords: user modeling, behavior model, forecasting, taxpayers, temporal series, neural networks.

RESUMO

Neste trabalho, propõe-se um modelo de usuário genérico, obtido a partir dos dados históricos obtidos das ações do usuário e de funções de inferência de comportamento. As funções de comportamento são obtidas através de técnicas de forecasting (previsão). O modelo em questão foi aplicado a um tipo especial de usuários, contribuintes de sistemas fiscais. Duas técnicas de previsão foram usadas: modelo estatístico de Holt e redes neurais artificiais. Os resultados obtidos na implementação deste modelo no Governo do Maranhão são mostrados

Palavras-chave: modelagem de usuários, modelagem comportamental, previsões, contribuintes fazendários, séries temporais, redes neurais.

ÍNDICE

1	INTRODUÇÃO	1
1.1	OBJETIVOS	2
1.2	Justificativa	3
1.3	Estrutura da Dissertação	3
2	REVISÃO BIBLIOGRÁFICA	5
2.1	Modelagem de Usuários	5
2.1.1	Conhecimento do usuário	6
2.1.2	Intenções do usuário	8
2.1.3	Preferências do Usuário	9
2.1.4	Aquisição e Representação de Modelos de Usuários	10
2.2	AGENTES INTELIGENTES E MODELOS DE AGENTES	11
2.3	FORECASTING (PREVISÕES)	14
2.3.1	<i>Elaboração de sistema de forecasting</i>	15
2.3.1.1	Definição do Problema	17
2.3.1.2	Coleta de Informações	19
2.3.1.3	Seleção do Pacote Computacional	21
2.3.1.4	Verificação do Sistema	24
2.3.2	<i>Modelos de previsão</i>	24
2.3.3	<i>Modelo de Holt</i>	25
2.3.4	<i>Modelos Autoregressivos – Média Móvel</i>	26
2.3.5	<i>Modelos de redes neurais</i>	27
2.4	TRABALHOS CORRELATOS	32
3	MODELO COMPORTAMENTAL DE CONTRIBUINTES	36
3.1	MODELO GENÉRICO	36
3.1.1	<i>Estado Interno</i>	37
3.1.2	<i>Percepção</i>	37
3.1.3	<i>Modelagem Comportamental</i>	37
3.1.4	<i>Modelo Computacional Proposto</i>	38
3.2	MICRO-MODELO DE CONTRIBUINTE	39
3.2.1	<i>Valor do imposto (V_i)</i>	40
3.2.2	<i>Probabilidade de ser descoberto (PSD)</i>	41
3.2.3	<i>Sanções (V_s)</i>	42
3.2.4	<i>Modelo Comportamental</i>	42
3.3	MACRO-MODELO DE CONTRIBUINTES	44
4	MODELO DE AGENTES PARA SOCIEDADE FAZENDÁRIA	46
4.1	ARQUITETURA	46
4.2	MODELAGEM DO AGENTE CONTRIBUINTE	49
4.2.1	<i>Conjunto modelo do contribuinte – M</i>	50
4.2.2	<i>Vetor de estado interno – V_{ei}</i>	50
4.2.3	<i>Pacote de percepções - P_{pc}</i>	51
4.2.4	<i>Funções de inferência – $V_{ei}(t)$</i>	51
4.2.5	<i>Comportamento do Agente</i>	52
5	ESTUDO DE CASO: SEFAZ-MA	53
5.1	IMPLEMENTAÇÃO	55
5.1.1	<i>Módulo estatístico</i>	56
5.1.2	<i>Módulo neural</i>	56
5.2	RESULTADOS OBTIDOS	59
5.2.1	<i>Módulo estatístico versus Módulo neural</i>	60

5.2.2	<i>Modelagem de grupos versus modelagem individual de contribuintes</i>	<i>61</i>
6	CONCLUSÕES	64
	REFERÊNCIAS BIBLIOGRÁFICAS	66
	ANEXO I – INTERFACE DO PACOTE COM CÓDIGO FONTE	71
	ANEXO II – RESULTADOS	74
	LEGENDA:	74
	SÉRIE EF HOLT	74
	SÉRIE EF NEURAL	74
	SÉRIE EC HOLT	74
	SÉRIE EC NEURAL	75
	SÉRIE SF HOLT	75
	SÉRIE SF NEURAL	75
	SÉRIE SC HOLT	76
	SÉRIE SC NEURAL	76
	SÉRIE VIF HOLT	76
	SÉRIE VIF NEURAL	77
	SÉRIE VIC HOLT	77
	SÉRIE VIC NEURAL	77
	SÉRIE PSD HOLT	78
	SÉRIE PSD NEURAL	78
	SÉRIE VS HOLT	78
	SÉRIE VS NEURAL	79

LISTA DE FIGURAS

Figura 1 - Aplicações para modelos de usuários	5
Figura 2 - Representações do usuário através de estereótipos.....	8
Figura 3 - Formas de aquisição de modelos de usuários	11
Figura 4 - Relação entre precisão e custo do <i>forecasting</i>	17
Figura 5 – Topologia genérica de RNA para previsão	28
Figura 6 - Topologia de RNA para previsão segundo Varfis e Versino.....	29
Figura 7 - Modelo genérico de usuário.....	37
Figura 8 - Arquitetura simplificada utilizada para simulação	47
Figura 9 - Arquitetura detalhada utilizada para simulação.....	49
Figura 10 - Modelo relacional para extração das séries comportamentais.....	53
Figura 11 - Arquitetura de implementação do sistema de previsão comportamental	55
Figura 12 - RNA utilizada no módulo neural de previsão.....	58

LISTA DE TABELAS

Tabela 1 – Atributos do modelo versus datawarehouse	54
Tabela 2 - Erro médio de Vic obtido (Holt versus RNA).....	60
Tabela 3 - Parâmetros da RNA utilizada.....	61
Tabela 4 - Erro médio de Vic obtido com modelo de Holt (grupal versus individual).....	62
Tabela 5 - Erro médio de Vic obtido com modelo de RNA (grupal versus individual).....	62

1 INTRODUÇÃO

A modelagem de usuários com o objetivo de conhecimento do seu comportamento recebe a atenção dos pesquisadores há vários anos [Perrault, 1978][Finin, 1986][Kobsa, 1989][Kobsa, 1993][Kobsa, 2001]. Os esforços concentram-se tanto nos aspectos de criação de um modelo adequado quanto à evolução desse modelo, que deve ocorrer em paralelo a alterações comportamentais do usuário. A Inteligência Artificial (IA), mais especificamente, os sistemas multiagentes, têm-se proposto a modelar processos comportamentais, indo desde modelos reativos [Bounabat, 2001] até cognitivos [Wooldridge 2002]. Nos modelos reativos utilizam-se agentes que não possuem um modelo de raciocínio e agem com base em respostas a estímulos externos. Por outro lado, os agentes cognitivos, ou deliberativos, são os que possuem um modelo simbólico de raciocínio e um plano a ser realizado e/ou negociado com outros agentes para alcançar seus objetivos.

Diversas são as aplicações para tal modelagem, destacando-se interação homem-computador, interfaces inteligentes, engenharia cognitiva, recuperação inteligente de informação, sistemas tutores inteligentes, sistemas especialistas e sistemas de simulação [Kobsa, 1993].

No âmbito tributário, a simulação de processos através de sistemas de simulação é reconhecidamente algo bastante eficaz para um bom planejamento tributário. Neste contexto, a administração fazendária brasileira, para fazer face às suas necessidades de investimento, vem estudando alternativas para aumentar a arrecadação de tributos. Uma vez que a origem de tais tributos advém dos contribuintes inseridos em cada esfera

tributária, o prévio conhecimento do comportamento desses contribuintes é um objeto extremamente facilitador para a melhoria do processo de planejamento. São inúmeras as aplicações que podem utilizar modelos comportamentais e econômicos. Entretanto, no caso das sociedades com grande concentração de renda, os tributos arrecadados tendem a vir de um pequeno número de contribuintes. Nestes casos, modelos de comportamento de grupo podem esconder alguns padrões individuais de comportamento. Isso traz à tona a necessidade de estudos que possam propor modelos comportamentais a serem aplicados individualmente nos contribuintes fiscais.

Este trabalho aborda o aspecto de modelagem individual do comportamento de usuários, em especial os contribuintes fazendários.

1.1 Objetivos

O objetivo principal é construir um modelo computacional que permita inferir dados de ações futuras do usuário em questão a partir de dados de suas ações passadas. O termo usuário é aplicado ao contribuinte devido a sua característica de interação com o fisco, sendo tal contribuinte considerado um usuário do sistema fiscal. As suas interações dizem respeito aos valores percebidos em suas variáveis econômicas, tais como entradas e saídas de insumos. Outro objetivo é desenvolver um protótipo para realização de simulações a partir dos modelos comportamentais individuais obtidos, avaliando-se os erros percebidos em relação ao comportamento real verificado por cada contribuinte.

Especificamente, pretende-se:

1. Propor um modelo comportamental de usuário baseado em dados de ações passadas;

2. Aplicar o modelo comportamental aos contribuintes fazendários;
3. Implementar o modelo na Secretaria de Estado da Fazenda do Maranhão.

1.2 Justificativa

No âmbito tributário, para facilitar o planejamento tributário e diminuir a evasão fiscal, alguns trabalhos têm dado uma ênfase em aspectos como classificação e seleção de contribuintes para fiscalização e modelos de gerenciamento de arrecadação utilizando técnicas estatísticas (cuja qualidade depende do tamanho da amostra em questão), como [Barreto, 2005][Corvalão, 2002][Liebel, 2004][Braz, 2001].

Entretanto, a realidade da maioria das administrações tributárias é que uma pequena minoria de contribuintes é responsável pela maior parte da movimentação econômica em uma esfera tributária. O estudo de grupos de contribuintes acaba-se tornando algo complexo, uma vez que a formação de tais grupos é uma tarefa por si só complexa e cujos resultados podem mascarar alguns padrões que só podem ser detectados individualmente. Desta forma, estudos de comportamento de contribuintes individualmente tornam-se atraentes e sendo assim, uma solução que possa ser utilizada, com eficiência, tanto em grandes quanto em pequenas amostras torna-se necessária.

1.3 Estrutura da Dissertação

No capítulo 2 é feita uma revisão bibliográfica com o objetivo de se obter um embasamento teórico para o trabalho, abordando-se os temas de modelagem de usuários e agentes inteligentes. Ainda neste capítulo, é abordado o assunto *forecasting*, descrevendo as principais técnicas de previsão utilizadas atualmente, a saber, os modelos estatísticos e os

modelos de redes neurais. São ainda comentados os principais trabalhos que buscam, de uma forma ou outra, a modelagem comportamental de sociedades fazendárias.

O capítulo 3 propõe um modelo para comportamento individual de contribuintes, através de variáveis relacionadas, concluindo com um macro-modelo de contribuintes, que permite que o modelo individual proposto possa ser expandido para a sociedade tributária como um todo.

No capítulo 4 é proposta uma arquitetura para uma sociedade de agentes, onde cada agente representa um contribuinte individualmente, detalhando-se a modelagem do ambiente de interações e do agente contribuinte, enquanto no capítulo 5 são mostrados os resultados obtidos ao se aplicar o modelo da sociedade proposta na Secretaria de Fazenda do Estado do Maranhão, com o Imposto sobre Circulação de Mercadorias e Serviços de Transporte – ICMS.

Finalmente, o capítulo 6 apresenta as conclusões a respeito da abordagem proposta e discute a continuidade do trabalho.

2 REVISÃO BIBLIOGRÁFICA

2.1 Modelagem de Usuários

Há quase três décadas a modelagem de usuários tem progredido amplamente. A necessidade de sistemas que se adaptassem automaticamente aos seus usuários foi um dos fatores que mais contribuiu para este desenvolvimento. Entretanto, com modelos de usuários criam-se oportunidades para se conseguir uma gama bem maior de aplicações que utilizem estes modelos, tais como interação homem-computador, interfaces inteligentes, recuperação inteligente de informação, sistemas tutores inteligentes, sistemas especialistas e sistemas de simulação [Kobsa, 1993], conforme ilustra a figura 1.

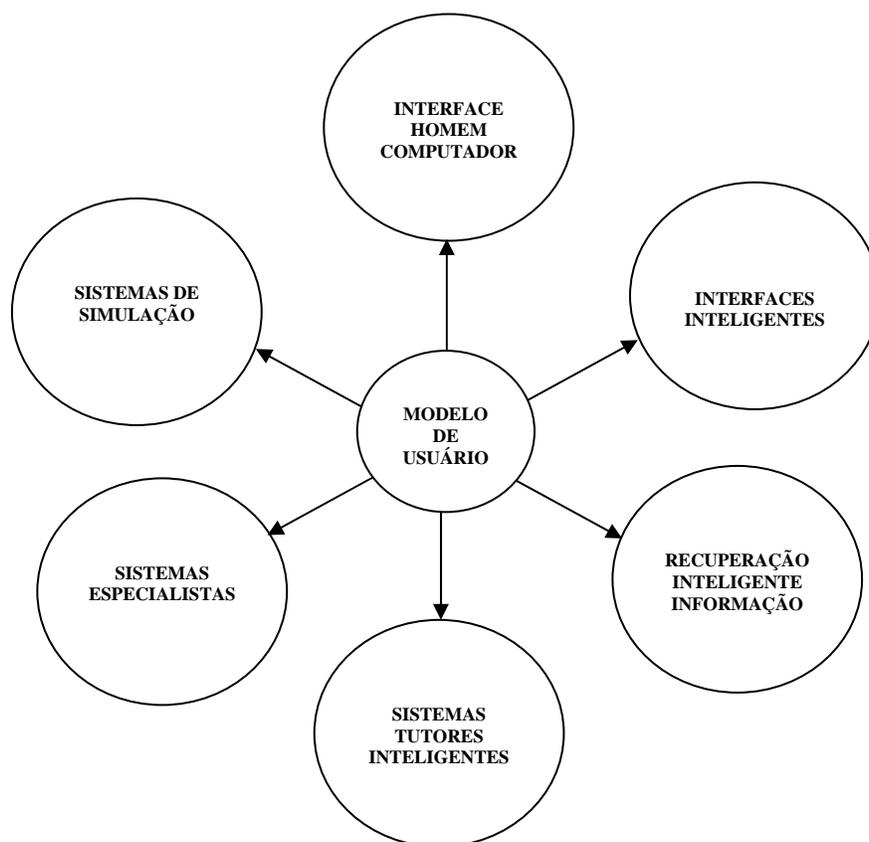


Figura 1 - Aplicações para modelos de usuários

O modelo de usuário é uma fonte de conhecimento que contém informações dos aspectos relevantes ao usuário. Um componente da modelagem de usuário formula suposições sobre o usuário baseado na interação com ele; armazena-as em uma estrutura de representação apropriada; infere novas suposições com base nas iniciais; mantém a consistência de um conjunto de suposições e fornece a outros componentes de um sistema essas suposições sobre o usuário. Portanto, através da mesma pode-se sintetizar as características e habilidades de usuários ou grupo de usuários com o objetivo de conhecer o seu comportamento.

Várias técnicas têm sido usadas para tentar sanar as dificuldades encontradas na modelagem de usuários. Estereótipos [Rich, 1979][Rich, 1983][Ballim, 1991][Chin, 1988], raciocínio padronizado [Jones, 1985][Csinger, 1989][Arragon, 1990] e análises de Dempster-Schafer [Shafer, 1976] são algumas dessas técnicas.

Dentre as principais tarefas dentro da abordagem da modelagem de usuários estão a necessidade de se obter uma aproximação do conhecimento, das intenções e das preferências do usuário, tratados a seguir.

2.1.1 Conhecimento do usuário

Os estereótipos [Rich, 1979][Rich, 1983][Ballim, 1991][Chin, 1988] tem sido amplamente utilizados para as áreas de aplicações onde um rápido, mas não necessariamente completo, conhecimento do usuário faz-se necessário. Nesta abordagem, o modelador de usuário tem basicamente três tarefas:

- Identificação do subgrupo do usuário: O modelador deve identificar subgrupos com a população de usuários cujos membros apresentam certa semelhança em relação a características relevantes à aplicação. Por exemplo, um banco de dados contendo informações fiscais de contribuintes pode ser dividido em contribuintes estáveis e instáveis, e esses grupos podem ainda ser subdivididos em sonegadores e não sonegadores.
- Identificação de características chaves: O modelador pode identificar um pequeno número de características que possibilitam classificar os membros de um subgrupo (a presença ou ausência dessas características deve ser detectada pelo sistema). No caso do banco de dados fiscal, tais características podem ser, por exemplo, o volume da movimentação financeira contribuinte ou então sua atividade econômica.
- Representação em estereótipos: As características relevantes à aplicação dos grupos de usuários identificados devem ser formalizadas em um sistema de representação apropriado. O conjunto de todas as características de um subgrupo de usuário é denominado estereótipo. Se o conteúdo de um estereótipo formar um subconjunto de um outro estereótipo, devem ser construídas hierarquias de estereótipos, onde um estereótipo subordinado a outro herda o seu comportamento. A figura 2 ilustra esta situação.

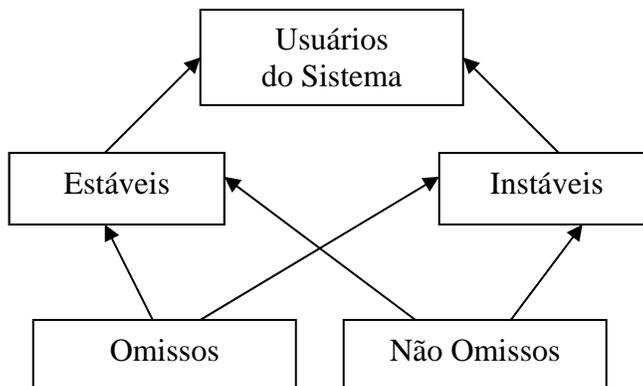


Figura 2 - Representações do usuário através de estereótipos

2.1.2 Intenções do usuário

A descoberta das intenções, também referenciada na literatura como planos de usuários, é algo bastante útil para a modelagem de usuários. Nestes planos, uma intenção é uma seqüência de ações do usuário que buscam um determinado objetivo [Raskutti & Zukerman, 1991][Appelt & Pollack, 1992][Mayfield, 1992][Eller & Carberry, 1992]. Sistemas de descoberta de planos observam as ações de entrada dos usuários e tentam determinar todos os planos do usuário para os quais as ações observadas podem ser completadas. Este conjunto de planos candidatos para o plano de usuário desconhecido pode ser rapidamente descartado à proporção que as ações do usuário mudam com o tempo.

Enquanto o algoritmo básico é extremamente simples, problemas de combinação surgem quando os mesmos são empregados na prática, principalmente devido a duas razões:

- Não há uma exatidão sobre em que momento uma se inicia um novo plano;

- Ações e curtas seqüências de ações podem fazer parte de mais de um plano;
- Os usuários podem interromper ou mesmo suspender a execução de seus planos correntes

Duas classes de técnicas têm sido empregadas para o reconhecimento de planos de usuários:

- Bibliotecas de planos: Nesta abordagem, todos os planos de usuários possíveis já são pré-armazenados em uma biblioteca. A seqüência de ações do usuário observada é comparada com estes planos pré-armazenados, e todos os planos cujo início coincide com a entrada do usuário são selecionados. As variações em um plano devem ser armazenadas como novos planos de usuários.
- Construção de planos: Nesta abordagem, o sistema possui uma biblioteca de todas as possíveis ações do usuário, inclusive com os efeitos e as pré-condições destas ações. A ação do usuário é completada por todas as seqüências de ações que preenchem os requisitos de que os efeitos das ações precedentes levam às pré-condições das ações subseqüentes.

2.1.3 Preferências do Usuário

Um considerável número de trabalhos envolvendo a modelagem das preferências dos usuários tem sido desenvolvido atualmente, principalmente no que se

refere às informações de suas preferências. Exemplos clássicos são os sistemas de recuperação de informação inteligentes [Brajnik, 1987] e filtragem de informação [Jennings, 1993].

2.1.4 Aquisição e Representação de Modelos de Usuários

A aquisição de modelos de usuários consiste exatamente em obter o conhecimento do usuário, suas intenções e suas preferências [Kobsa, 1993]. Dessa forma, essa é a maior complexidade para que se adquiram modelos utilizáveis para resolver algum problema.

A aquisição dos modelos pode ser feita basicamente de três formas distintas, conforme ilustra a figura 3:

- Explícita: adequado quando os usuários têm interesse em serem modelados ou quando a massa de informações a serem analisadas é considerada de pequena a média. Aquisições explícitas podem ser feitas através de entrevistas e questionários.
- Implícita: adequado quando os usuários não têm interesse em serem modelados ou quando a massa de informações a serem analisadas é considerada grande. Podem ser feitas através de observação direta ou análise de fatos.
- Híbrida: adequada em ambientes híbridos, onde existem usuários que ora querem e ora não querem que seu comportamento seja detectado ou quando a massa de informações disponíveis varia entre usuários distintos.

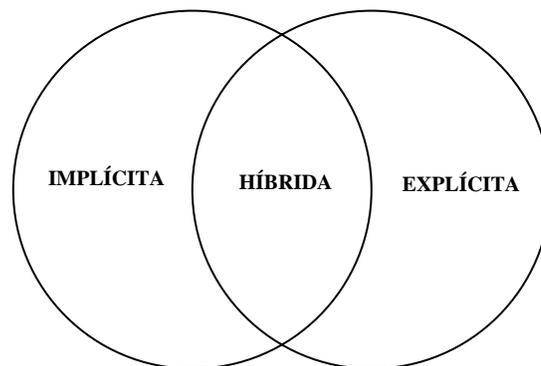


Figura 3 - Formas de aquisição de modelos de usuários

Além disso, algo também muito importante é a escolha da forma de representação do conhecimento do usuário. Problemas distintos podem levar a formas de representação distintas, pois se deve escolher aquela que melhor se adapte ao problema em questão. Exemplos de formas de representação são ontologias (redes conceituais) para representação de vocabulários dos usuários ou vetores para representar seus estados internos.

2.2 Agentes Inteligentes e Modelos de Agentes

A abordagem descentralizada tem sido muito utilizada para a simulação de sistemas complexos. Esta abordagem propõe que o comportamento complexo de alguns sistemas é decorrente das interações entre seus componentes. A tendência para assumir a existência de um controle central, subjacente aos sistemas, torna difícil o entendimento de muitos fenômenos ao nosso redor. Em vez de utilizar macro-modelos, para reproduzir o comportamento observável dos sistemas, a abordagem descentralizada introduz o conceito de micro-simulação. Este método consiste em construir modelos microscópicos com a

capacidade de reproduzir as leis macroscópicas do sistema, através da simulação de cada componente individualmente, quando interagem entre si [Claramunt & Jiang, 2001].

[Wooldridge, 1994] define um agente como uma entidade autônoma, que possui um sistema interno de tomada de decisões, agindo sobre o mundo e sobre outros agentes que o rodeiam e por fim, que é capaz de funcionar sem a necessidade de algo ou alguém para guiá-lo. De posse de bases de conhecimento e de mecanismos de raciocínio, os agentes devem ser capazes de reconhecer situações em que devem se ativar, sem que o usuário perceba, ou seja, de forma transparente ao usuário.

Ainda segundo [Wooldridge, 1994], a construção e a especificação da estrutura e funcionamento de um agente genérico pode ser realizada segundo três tipos de arquiteturas:

- Arquiteturas deliberativas: segue a abordagem clássica da Inteligência Artificial, onde os agentes contêm um modelo simbólico do mundo, explicitamente representado, e cujas decisões (ações) são tomadas via raciocínio lógico, baseadas em casamento de padrões e manipulações simbólicas;
- Arquiteturas reativas: a arquitetura reativa é aquela que não inclui nenhum tipo de modelo central e simbólico do mundo e não utiliza raciocínio complexo e simbólico. Baseia-se na proposta de que um agente pode desenvolver inteligência a partir de interações com seu ambiente, não necessitando de um modelo pré-estabelecido;

- Arquiteturas híbridas: a arquitetura híbrida mistura componentes das arquiteturas deliberativas e reativas com o objetivo de torná-la mais adequada e funcional para a construção de agentes.

Essas considerações feitas sobre agentes levam a crer que a modelagem de usuários pode ser facilitada se for implementada com agentes inteligentes, além da melhora na eficiência e eficácia do sistema, através dos mecanismos de adaptação.

Um modelo baseado em agentes consiste na criação de uma população de agentes, com capacidade de percepção e ação similar aos componentes simulados. Para que os agentes possam agir, como se fossem os próprios componentes de um sistema, deve-se dotá-los de comportamentos (regras que definem suas possíveis ações). Isto é feito através da modelagem de comportamentos, que consiste em analisar os componentes de um sistema e extrair as suas principais características para que possam ser incorporadas nos respectivos agentes. Desta forma, a modelagem de comportamentos é uma etapa fundamental na criação de um modelo baseado em agentes.

Utilizando a abordagem descentralizada, é possível simular o comportamento de usuários, a partir de um modelo que seja capaz de reproduzir o comportamento de uma só pessoa (modelo individual). Segundo esta abordagem, os padrões (ou efeitos) observados na forma como as pessoas comportam-se, podem ser entendidos como o resultado das decisões tomadas por cada usuário em decorrência da sua necessidade de encontrar uma situação agradável. Assim, a compreensão do indivíduo em relação ao ambiente no qual está inserido se torna muito importante para entender as formas de se atingir o objetivo desejado.

De forma geral, para que os agentes possam realizar interações entre si, de forma objetiva, faz-se necessário dotá-los de mecanismos para perceber o ambiente e para realizar ações (especificamente interações com o fisco). Além destes mecanismos, o modelo deve permitir que os agentes possam ter algumas propriedades para indicar seu estado atual (por exemplo, a satisfação e o grau de interação com o ambiente) na simulação.

2.3 *Forecasting* (Previsões)

As previsões desempenham um importante papel em diversas áreas na gestão de organizações [Makridakis, 1998][MENTZER, 1997]; por exemplo, na área financeira (no planejamento da necessidade de recursos), na área de recursos humanos (no planejamento de modificações no nível da força de trabalho), na área de vendas (no agendamento de promoções) e na área tributária (planejamento tributário).

Previsões são elaboradas utilizando métodos quantitativos, qualitativos ou combinações de ambos. Métodos quantitativos, neste trabalho denominados métodos de *forecasting*, baseiam-se na análise de séries temporais (dados que descrevem a variação dos dados ao longo do tempo). Métodos qualitativos baseiam-se em opiniões de especialistas, sendo vulneráveis a tendências que podem comprometer a confiabilidade de seus resultados. Métodos qualitativos têm sido, historicamente, os mais utilizados na previsão de dados [Mentzer & Cox, 1984]. Tais métodos costumam apresentar um baixo grau de precisão; apesar disto, continuam sendo amplamente utilizados nas organizações, mesmo com a difusão dos métodos de *forecasting* mais avançados, impulsionada pelo avanço na capacidade de processamento e armazenamento de dados computacionais [Sanders & Manrodt, 1994]. A extensa utilização dos métodos qualitativos parece estar relacionada ao

fato das previsões por eles geradas corresponderem às metas de demanda estabelecidas pelas organizações [Dias, 1999]. A escassa fundamentação prática dessas previsões, por outro lado, pode explicar, em grande parte, a baixa precisão dos resultados oferecidos pelos métodos qualitativos.

A tomada de decisões é um fato cotidiano que desempenha um papel relevante dentro das organizações. Atualmente, o alto grau de competitividade no meio empresarial exige a capacidade de tomar decisões rápidas e precisas. A qualidade da tomada de decisão tem relação direta com os dados disponíveis para o tomador de decisão e com sua habilidade em extrair destes dados informações relevantes.

Através das técnicas de *forecasting* é possível extrair dos dados passados disponíveis sobre um processo, informações que permitem a modelagem matemática de seu comportamento. A suposição de uma continuidade nesse comportamento permite a realização de previsões, cuja qualidade e precisão são muito superiores àquelas das previsões feitas intuitivamente, baseadas unicamente na experiência dos decisores.

Adicionalmente, os modelos uma vez atualizados, passam, de imediato, a refletir as alterações do processo, fornecendo prontamente subsídios a novas tomadas de decisão.

São inúmeras as aplicações de *forecasting* dentro de uma organização. A operacionalização satisfatória de estratégias de planejamento está fortemente associada à existência de um sistema eficiente de *forecasting* [Hill, 1994; Tompkins et al., 1996].

2.3.1 Elaboração de sistema de *forecasting*

A elaboração de um sistema de *forecasting* requer, de uma organização, conhecimento e habilidade em 4 áreas básicas: (i) identificação e definição dos problemas a serem tratados no *forecasting*; (ii) aplicação dos métodos de *forecasting*; (iii) procedimentos para seleção do método apropriado a situações específicas; e (iv) suporte organizacional para adaptar e usar os métodos de *forecasting* requeridos.

Um sistema de *forecasting* deve estabelecer relações entre previsões feitas pelas diferentes áreas de gerenciamento. Existe um alto grau de dependência entre essas previsões; o perfeito entendimento da natureza desta dependência define o sucesso na implantação do sistema de *forecasting*. Por exemplo, um erro na projeção de vendas pode trazer uma série de prejuízos a todas as demais áreas da organização.

A aplicabilidade de um sistema de *forecasting* depende de três condições [Makridakis et al., 1998]:

- Disponibilidade de informações históricas;
- Possibilidade da transformação das informações históricas em dados numéricos; e
- Suposição da repetição de padrões observados em dados passados no tempo futuro.

Esta última consideração é conhecida como suposição de continuidade. Tal condição é uma premissa básica para a utilização de métodos de *forecasting*, bem como de diversos métodos qualitativos.

As técnicas de *forecasting* variam consideravelmente, tendo sido desenvolvidas com vários propósitos distintos. Cada técnica possui características próprias, grau de precisão e custo de utilização, os quais devem ser considerados na escolha de um método específico. De forma geral, os critérios para a escolha de um sistema de *forecasting* compreendem as etapas descritas na seqüência.

2.3.1.1 Definição do Problema

Em algumas aplicações de *forecasting*, a definição do problema pode ser a etapa mais complexa. Diversos fatores devem ser analisados: como o *forecasting* será usado, onde será usado e como ele se encaixa dentro da organização. O nível de detalhe requerido é uma consideração de extrema importância, sendo influenciado por diversos fatores, tais como disponibilidade de dados, precisão, custo da análise e preferências gerenciais.

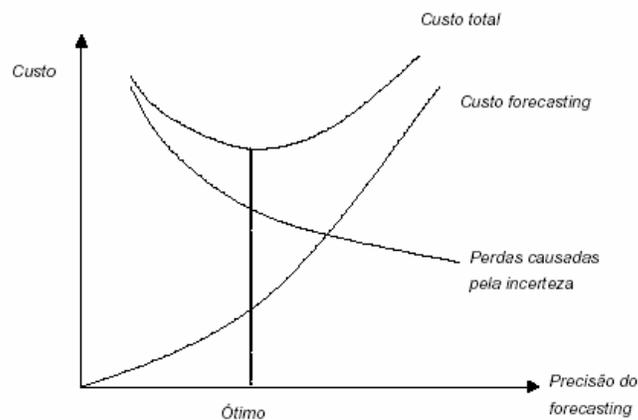


Figura 1. Relação entre precisão e custo do *forecasting*. (Adaptado de Montgomery *et al.*, 1990).

Figura 4 - Relação entre precisão e custo do *forecasting*

O custo do *forecasting* está diretamente ligado à precisão requerida. Uma vez que o aumento da precisão diminui as perdas resultantes dos processos decisórios, a relação

entre o custo do *forecasting* e as perdas causadas pela incerteza forma um *trade-off*, exemplificado na figura 4.

Fica claro, a partir da análise do gráfico desta figura, que após um determinado ponto o aumento dos recursos investidos não implica em aumento expressivo na precisão do sistema de *forecasting*. Sendo assim, procura-se trabalhar dentro de uma faixa que possibilite a melhor previsão a um menor custo.

Uma segunda classe de decisões envolve elementos temporais; mais especificamente, o período, horizonte e intervalo do *forecasting*.

O período é a unidade básica de tempo em que a previsão é requerida. Geralmente, ele é expresso em meses ou semanas, dependendo do espaço de tempo em que os dados de demanda estão armazenados.

Dentre os elementos temporais, a magnitude do período é o fator que mais influencia na escolha do modelo a ser utilizado.

O horizonte é o número de períodos futuros cobertos pela previsão, sendo expresso na mesma unidade temporal do período. Ele está relacionado com a capacidade de resposta da organização. Quanto menos flexível for a organização, maior será o horizonte; quanto mais ágil, menor o horizonte. Recomenda-se que o horizonte do *forecasting* deva ser, no mínimo, igual ao maior tempo de resposta da organização [Dias, 1999].

O intervalo é a frequência com a qual as novas previsões são preparadas. Na definição desta frequência, existe um *trade-off* entre o risco de não se identificar uma

mudança na série temporal e os custos na revisão do *forecasting*. Assim, o intervalo depende da estabilidade do processo, das conseqüências de se estar usando uma previsão obsoleta, e dos custos do *forecasting* e do re-planejamento. Geralmente o intervalo é igual ao período.

Desta maneira, os modelos são revistos a cada período, usando a demanda do período mais recente [Montgomery et al., 1990].

Durante a etapa de definição do problema, o técnico responsável pela elaboração do sistema de *forecasting* deve consultar todos aqueles envolvidos na coleta de dados, na manutenção do banco de dados e no uso das previsões para o planejamento futuro.

2.3.1.2 Coleta de Informações

Pelo menos dois tipos de informações devem estar disponíveis na elaboração de um sistema de *forecasting* [Makridakis et al., 1998]: (i) dados estatísticos (geralmente numéricos) e (ii) dados subjetivos oriundos de julgamento e perícia de especialistas (empregado, principalmente, na avaliação da qualidade dos dados a serem utilizados no sistema). Os dados estatísticos serão utilizados na modelagem matemática do *forecasting*; a opinião de especialistas é essencial para validação prática das previsões geradas pelo sistema.

a Montagem do Banco de Dados

Os dados estatísticos a serem utilizados na previsão do processo são usualmente armazenados em um banco de dados. O banco de dados deve conter, além da série temporal (representada por elementos e uma quantificação dos mesmos em cada período),

informações que possibilitem a utilização de filtros. O filtro, em um banco de dados, é a designação utilizada para quaisquer critérios empregados no agrupamento de dados. Considere, por exemplo, a estratificação da previsão de um elemento específico, onde esse elemento faz parte de um conjunto maior e contém vários conjuntos menores; os filtros, no exemplo, são o superconjunto, ou um subconjunto específico.

A atualização do banco de dados deve ser feita a cada período, incorporando-se, assim, as informações mais recentes aos modelos de previsão.

b Definição dos Níveis de Agregação

Algumas vezes, para se captar alguma característica no comportamento de uma série temporal, faz-se necessário uma agregação de seus elementos temporais. Por exemplo, quando se está analisando uma série contendo dados diários, pode não ser possível visualizar com facilidade um padrão no comportamento da série que se manifeste numa base mensal, o que poderia implicar em uma deficiência na modelagem. Neste caso, poder-se-ia agregar os dados em períodos semanais e mensais, até que algum padrão possa ser identificado na série de maneira distinta.

Muitas vezes, mesmo após várias agregações temporais, nenhuma característica no comportamento da série é encontrada. Isto significa que os dados de demanda apresentam um comportamento exclusivamente aleatório, o que pode comprometer a precisão de previsões futuras.

As agregações temporais são definidas, basicamente, pelos dados de demanda disponíveis. Já as agregações dos produtos são definidas por preferências gerenciais, custos

da análise, precisão requerida e a disponibilidade dos dados de demanda [Montgomery et al.,1990].

2.3.1.3 Seleção do Pacote Computacional

Dada a complexidade de operacionalização de alguns dos modelos de *forecasting*, faz-se necessário o uso de pacotes computacionais no cálculo da previsão. A escolha correta do pacote adequado pode ser difícil, devido a grande variedade de produtos disponíveis no mercado. A seguir, são apresentadas algumas questões que podem ser úteis na determinação do pacote computacional mais apropriado para apoio a um sistema de previsão de demanda; a lista apresentada foi elaborada a partir de sugestões em [Makridakis et al., 1998]:

- O pacote deve possuir vantagens identificadas como essenciais. Desta forma, deve-se verificar os modelos de *forecasting* contemplados no produto, a forma de gerenciamento das informações, a apresentação gráfica e relatórios dos resultados obtidos na análise;
- Identificar o sistema operacional do pacote. O sistema deve ser compatível com aquele utilizado pelos computadores da organização, ou permitir a transferência de dados entre sistemas operacionais distintos;
- O pacote deve ser de fácil utilização e aprendizado. Desta forma, deve-se solicitar uma demonstração de funcionamento do programa que verifique aspectos relacionados a sua facilidade de operacionalização;

- Verificar a possibilidade de implementação de novos modelos de *forecasting* no pacote computacional. Usuários avançados procuram fazer modificações em modelos existentes de *forecasting* ou mesmo implementar novos modelos nos pacotes. Para tanto, a linguagem de programação do pacote selecionado deve ser dominada pelos usuários (programadores) da organização;
- Muitas vezes, centenas de séries temporais podem estar em estudo. Alguns pacotes possuem uma ferramenta que permite gerar rapidamente a previsão individual de um conjunto de dados envolvendo milhares de séries temporais. Esta característica é de suma importância quando se necessita agilidade na análise de muitas séries;
- Verificar a capacidade de processamento de dados do pacote. Sistemas de *forecasting* podem utilizar séries temporais bastante extensas, que eventualmente ultrapassam o limite de capacidade de processamento de alguns pacotes computacionais;
- Verificar a precisão das previsões calculadas pelo pacote. Apesar de possuírem diferentes algoritmos, pacotes distintos devem apresentar resultados no mínimo próximos. Assim, é interessante fazer uma comparação entre eles, uma vez que alguns podem conter erros.

A maioria dos pacotes estatísticos genéricos apresenta uma opção de análise de *forecasting*. Tais pacotes são recomendados a usuários que demandem uma análise estatística e gráfica detalhada dos dados. Entre os pacotes estatísticos genéricos mais

difundidos, três são abordados na sequência: NCSS, Statgraphics e SPSS. A análise destes pacotes limita-se aos seus módulos de *forecasting*.

O NCSS (*Number Cruncher Statistical System*, 1996) possui, em seu módulo de *forecasting*, a possibilidade de uso dos modelos ARIMA (tratado posteriormente), suavização exponencial e decomposição. O relatório de resultados da análise traz um grande número de informações, não disponíveis nos demais pacotes aqui abordados. O Statgraphics (1995) também possui as opções dos modelos de decomposição, ARIMA e suavização exponencial no seu módulo de *forecasting*, apresentando uma melhor interface de utilização, se comparado ao NCSS. O SPSS (*Statistical Package for Social Science*, 1997) é o melhor dentre os três pacotes genéricos aqui abordados. O programa oferece, no seu módulo de *forecasting*, as opções dos modelos de decomposição, ARIMA, suavização exponencial, regressão com erros do modelo ARIMA.

Os pacotes computacionais específicos para análise de *forecasting* contêm apenas as ferramentas estatísticas necessárias para a análise de séries temporais e geração de previsões. Sua vantagem se dá em funções relativas à previsão, não encontradas em pacotes de uso genérico. Entre os pacotes específicos para análise de *forecasting*, destacam-se o Forecast Pro e o Autobox.

O Forecast Pro (1999) oferece a opção de escolha automática do modelo que seja mais adequado a uma série temporal. Para tanto, faz uma comparação entre os modelos ARIMA, suavização exponencial e X-11 ARIMA. O pacote também permite ao usuário a seleção do critério utilizado para comparação dos modelos obtidos em cada série. Uma outra característica importante do pacote é a possibilidade de fazer, simultaneamente, a

previsão de um grande número de séries temporais em poucos segundos. O Autobox (1999) oferece, basicamente, as mesmas opções do pacote computacional Forecast Pro. Porém, utiliza os modelos ARIMA, suavização exponencial, além da análise de regressão a partir dos erros oriundos dos modelos ARIMA.

2.3.1.4 Verificação do Sistema

Uma vez tendo-se os modelos e seus parâmetros estimados apropriadamente, sua utilização na predição dos dados futuros do processo pode ser testada. Neste ponto, o processo de implantação do sistema de *forecasting* é considerado concluído, tendo início o seu processo de manutenção. A manutenção do sistema consiste na incorporação de novas informações sobre as variáveis de interesse, obtidas após cada período, e revalidação dos modelos estatísticos inicialmente selecionados para sua previsão futura.

2.3.2 Modelos de previsão

O emprego de cada modelo de previsão depende basicamente do comportamento da série temporal que se deseja analisar. As duas principais classes de modelos de previsão são os métodos estatísticos e os métodos de redes neurais. Dentre os métodos estatísticos, dois serão analisados: um de simples implementação, o modelo de Holt, e um dos mais utilizados para previsões em séries temporais, o ARIMA. Entretanto, para ambas as classes de métodos, algumas características em comum devem ser consideradas.

Uma série temporal pode exibir até quatro características diferentes em seu comportamento: média, sazonalidade, ciclo e tendência. A característica de média existe quando os valores da série flutuam em torno de uma média constante. A série possui

característica sazonal quando padrões cíclicos de variação se repetem em intervalos relativamente constantes de tempo. A característica cíclica existe quando a série exhibe variações ascendentes e descendentes, porém, em intervalos não regulares de tempo. Finalmente, a característica de tendência ocorre quando a série apresenta comportamento ascendente ou descendente por um longo período de tempo. Toda variação em uma série temporal que não pode ser explicada pelas características acima, é devida ao ruído aleatório no processo gerador dos dados; tal ruído não é matematicamente modelável.

A seguir são apresentados os principais modelos das famílias de modelos matemáticos de previsão comentados anteriormente.

2.3.3 Modelo de Holt

Uma forma bastante atraente devido a sua simplicidade e aos bons resultados obtidos em previsões de séries históricas é o modelo de Holt. Este modelo pertence a uma família de modelos denominada Modelos de Suavização Exponencial. Os modelos de suavização exponencial são amplamente utilizados para previsão de demanda devido a sua simplicidade, facilidade de ajuste e boa precisão. Estes métodos usam uma ponderação distinta para cada valor observado na série temporal, de modo que valores mais recentes recebam pesos maiores. Assim, os pesos formam um conjunto que decai exponencialmente a partir de valores mais recentes.

O modelo de Holt emprega duas constantes de suavização, α e β (com valores entre 0 e 1), sendo representado por três equações [Armstrong, 1999]:

$$L_t = \alpha z_t + (1-\alpha)(L_{t-1} + T_{t-1}) \quad (1)$$

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \quad (2)$$

$$Z'_{t+k} = L_t + kT_t. \quad (3)$$

As equações (1) e (2) fazem uma estimativa do nível e da inclinação da série temporal, respectivamente. Já a equação (3) calcula a previsão de valores para os próximos k períodos.

Nestas equações, L_t e T_t correspondem, nesta ordem, ao nível e à inclinação da série temporal no instante t , enquanto L_{t-1} e T_{t-1} referem-se ao nível e à inclinação da série no instante $t-1$; Z'_{t+k} corresponde ao valor da previsão feita para k períodos após o instante t ;

2.3.4 Modelos Autoregressivos – Média Móvel

Os modelos autoregressivos integrados a média móvel (ARIMA – Autoregressive Integrated Moving Average), também conhecidos como modelos de Box-Jenkins, foram propostos por George Box e Gwilym Jenkins no início dos anos 70 (Box et al., 1994).

Os modelos de Box-Jenkins partem da idéia de que os valores de uma série temporal são altamente dependentes, ou seja, cada valor pode ser explicado por valores prévios da série. Os modelos ARIMA representam a classe mais geral de modelos para a análise de séries temporais.

Um modelo estocástico útil na representação de um grande número de séries temporais é o modelo autoregressivo. Neste modelo, o valor corrente do processo é

expresso como uma combinação linear finita de valores prévios do processo e um ruído aleatório.

Entretanto, os modelos de Box-Jenkins têm uma hipótese básica de utilização referente ao tamanho da amostra. De acordo com (Box & Jenkins, 1976), este tamanho deve ser, no mínimo, de 50 observações. Este tende a ser um obstáculo significativo para muitos problemas cujos dados somente são coletados mensalmente ou mesmo anualmente.

2.3.5 Modelos de redes neurais

Os Modelos de Redes Neurais Artificiais (RNA) diferenciam-se dos modelos tradicionais de previsão por serem modelos não paramétricos envolvendo algoritmos de aprendizado [Abelém, 1994]. Tais algoritmos buscam imitar a estrutura de interconexões do cérebro humano, incorporando o padrão de comportamento de uma série temporal e prevendo valores futuros da mesma [Turban, 1993].

A construção de um Modelo RNA envolve desde a modelagem adequada da rede até as transformações utilizadas para transmitir os dados à mesma e métodos utilizados para interpretar os resultados obtidos. Esses três aspectos (modelagem, transformações e interpretação) são fundamentais na utilização das RNA para previsão.

O diagrama a seguir mostra uma rede neural com três camadas utilizada para soluções de problemas de previsão.

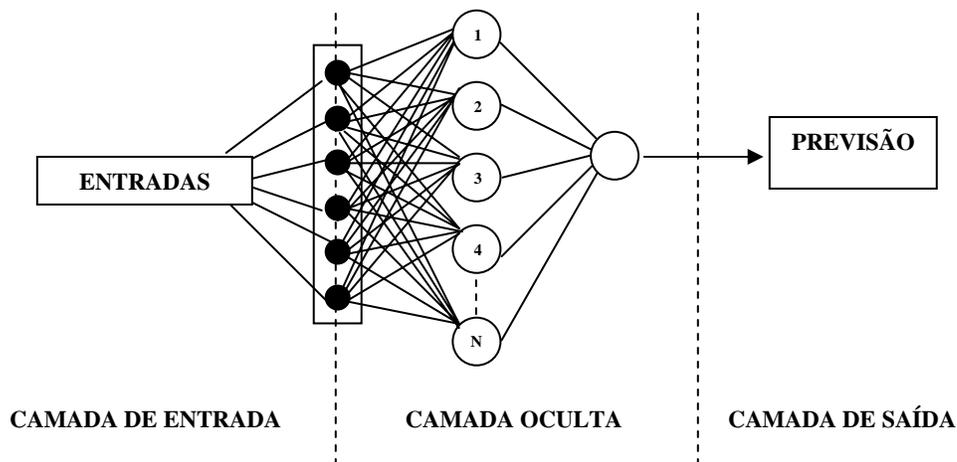


Figura 5 – Topologia genérica de RNA para previsão

Na camada de entrada são dispostas as variáveis que se julga serem úteis para representar o caso em questão. Em [Varfis e Versino, 1990], que trata da previsão de séries econômicas mensais, encontra-se uma proposta de metodologia onde procura-se modelar a entrada da rede de maneira a captar as componentes da série temporal. Desta forma as unidades foram divididas em três grupos:

- valores passados da série;
- valores passados referentes ao mesmo mês em anos anteriores, visando captar tendências ou ciclos da série;
- grupo de doze unidades binárias para captar a sazonalidade da série, tais como 100000000000 para o mês de janeiro, 010000000000 para fevereiro e assim sucessivamente, até 000000000001 para o mês de dezembro.

A figura a seguir apresenta de forma esquemática a topologia de uma RNA utilizando-se a abordagem de [Varfis e Versino, 1990].

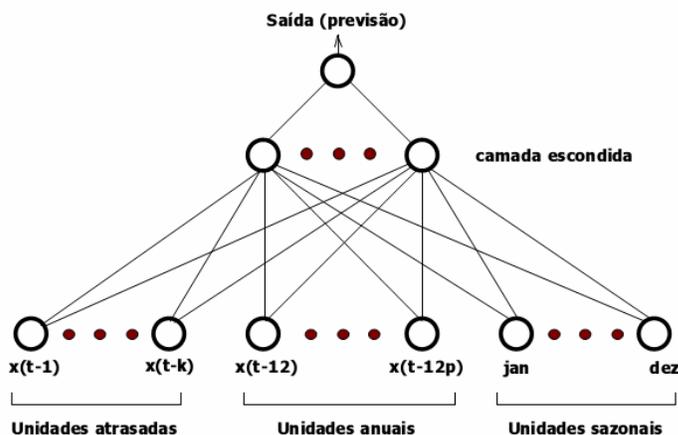


Figura 6 - Topologia de RNA para previsão segundo Varfis e Versino

Na camada intermediária, existem algumas regras clássicas utilizadas na definição da quantidade de elementos a serem usados. Nesta determinação é necessário evitar-se o under-learning (a rede não alcança um nível de desempenho aceitável) ou o over-learning (o erro começa a crescer). Neste último caso, deve-se buscar diminuir o número de neurônios ou mesmo o número de camadas ocultas, se possível.

Uma rede do tipo MLP (Multi Layer Perceptron) com duas camadas ocultas é teoricamente suficiente para modelar qualquer tipo de problema. Entretanto, na prática, redes com uma única camada oculta são as mais usuais. A quantidade de neurônios a serem utilizados na camada oculta é uma questão ainda nebulosa.

Muitos autores, como [Caudill,1990], apresentam o número de unidades das camadas escondidas como incógnita e simplesmente afirmam que "o tamanho da camada

escondida é escolhido usando o bom senso: se a camada for muito grande, a rede estará memorizando os padrões e com isso perdendo a capacidade de generalização, por outro lado, se a camada for muito pequena, a rede levará muito mais iterações para atingir a precisão desejada".

Como ponto de partida, uma boa referência prática é utilizar-se uma única camada com uma quantidade de elementos igual à metade da soma do número de unidades de entrada e de saída. Um parâmetro também muito citado em diversos trabalhos é o baseado no pressuposto de Hecht-Nielsen/Kolmogorov: “Uma rede neural com três camadas (entrada, oculta e saída) pode modelar qualquer função matemática contínua, desde que a camada oculta contenha $(2 I + 1)$ neurônios, onde I é o número de variáveis de entrada”.

Na camada de saída localizam-se os neurônios cujos axônios representam os valores de solução.

O algoritmo mais estudado e bem sucedido no processo de aprendizagem de uma RNA é o algoritmo de *backpropagation*. Através do algoritmo *backpropagation* [Rumelhart, 1986], são realizadas as transformações necessárias para a transmissão dos dados para a rede, possibilitando a formação de previsões consistentes com o comportamento da série em estudo.

Para o aprendizado, a amostra é dividida em duas faixas. Na primeira, denominada faixa-treino, a rede capta o padrão teste, testando se o erro é minimizado. A faixa treino deve ser suficientemente grande para captar o padrão de comportamento da

série, e deve ser representativa das condições futuras que a rede vai tentar captar para a realização da previsão.

O algoritmo de *backpropagation* ajusta então as ponderações, de modo a minimizar os erros de previsão dentro da faixa teste. Baseado nesse erro entre observado e estimado, os pesos são ajustados de modo a obter um erro mínimo.

Uma formalização do algoritmo *backpropagation* para previsão de séries de tempo pode ser encontrada nas seguintes equações, de acordo com [Corrêa e Portugal, 1998]:

$$\begin{aligned}\Delta w_{ij,k}^n &= \eta \cdot \delta_{j,k} \cdot y_{i,k-1} \\ w_{ij,k}^{n+1} &= w_{ij,k}^n + \Delta w_{ij,k}^n \\ \delta_{j,k} &= \left(\frac{dy}{dz} \right)_{i,k} \cdot (y_{i,k} - y_{i,k}^T)\end{aligned}$$

com: $i = 1, 2, \dots, I$; $j = 1, 2, \dots, J$, e $k = 1, \dots, K$.

onde: η é o coeficiente de aprendizado ($0,01 \leq \eta \leq 1,00$); $y_{i,k}$ é a saída do neurônio i na camada k ; $y_{i,k}^T$ é o valor desejado para $y_{i,k}$; $w_{ij,k}^n$ é a ponderação ligando o neurônio i na camada $k-1$ ao neurônio j na camada k no passo n .

Ao contrário dos demais modelos abordados anteriormente, os modelos de redes neurais artificiais não têm um procedimento universal de modelagem ou construção [Chatfield,1996].

2.4 Trabalhos Correlatos

A administração fazendária brasileira, para fazer face às suas necessidades de investimento, vem estudando alternativas para aumentar a arrecadação de tributos. Nesta linha, esforços estão sendo concentrados para coibir a evasão de receita e divulgar junto a comunidade empresarial a função social do imposto, sua relevância no contexto administrativo, seus reflexos na construção do Estado e na qualidade de vida de seus habitantes.

Os diferentes órgãos fazendários, responsáveis pelas administrações fazendárias tanto a nível municipal, estadual e federal, encontram-se hoje em diferentes estágios na busca de alternativas de melhoria de suas gestões administrativo-tributárias. No entanto, os principais problemas enfrentados por elas, variando em intensidade, dizem respeito às mesmas questões, tais como planejamento da ação fiscal, inteligência fiscal e previsão, acompanhamento e análise de receita.

Uma vez que a origem dos tributos advém dos contribuintes inseridos em cada esfera tributária, o prévio conhecimento dos seus perfis é um objeto extremamente facilitador na resolução dos problemas citados.

Alguns trabalhos têm dado uma ênfase em aspectos tais como classificação e seleção de contribuintes para fiscalização e modelos de gerenciamento de arrecadação utilizando técnicas estatísticas (cuja qualidade depende do tamanho da amostra em questão).

[Braz, 2001], analisa o problema da tributação no Brasil, e apresenta um modelo que permite o gerenciamento, avaliação e planejamento da arrecadação dos tributos de competência estadual, dando ênfase ao ICMS – Imposto sobre operações relativas à circulação de mercadorias e sobre prestações de serviços de transporte interestadual e intermunicipal e de comunicação, por ser este o tributo mais significativo para os cofres do governo. O modelo permite uma análise comportamental de atividades econômicas, descendo até o nível de faixas de faturamento por atividade, cruzando variáveis como faturamento, icms declarado, icms pago, débitos, créditos e energia elétrica para os conjuntos em questão. Através dessa análise de comportamento dos conjuntos, é possível chegar-se a indícios de evasão fiscal.

[Corvalão, 2002] tem como objetivo central verificar a possibilidade de melhorar as previsões mensais do Imposto sobre Circulação de Mercadorias e Serviços (ICMS) a ser arrecadado pelo Estado de Santa Catarina. Para tanto, utiliza um modelo comportamental global (modela o comportamento da arrecadação do grupo formado por todos contribuintes) de regressão dinâmica, baseando-se nos conceitos de co-integração e modelos de correção de erros e empregando a abordagem geral para específico sugerida pela *London School of Economics* (LSE).

[Arraes e Chumvichitra, 1996] fizeram uma aplicação de modelos comportamentais unidimensionais autoregressivos para o Estado do Ceará, utilizando dados trimestrais, obtendo um erro percentual absoluto médio de 3,5%;

[Ferreira, 1996], em sua dissertação de mestrado, emprega além de modelos unidimensionais, função de transferência para a previsão do ICMS no Estado do Ceará no período de 1970 a 1995. Apresentou um erro percentual absoluto médio de 4,8%;

[Coccaro, 2000], faz comparação das previsões usando redes neurais, modelo estrutural e modelo unidimensional ARIMA no Estado do Rio Grande do Sul, usando dados mensais de janeiro de 1981 a junho de 1999. Na melhor situação, obteve um erro percentual absoluto médio de 3,85% com redes neurais;

[Silveira, 2000], utilizou modelo unidimensional ARIMA na previsão do ICMS em Goiás, valendo-se de dados mensais de janeiro de 1995 a dezembro de 1999. Apresentou previsões fora da amostra para um período de três meses nos setores ação fiscal (EPAM = 5,65%) e dívida ativa (EPAM = 25,36%);

[Bernardo, 2001] empregou técnicas econométricas para previsão do ICMS no Estado de Roraima, não reportou o critério de acurácia para as previsões efetuadas.

[Liebel, 2004] aplica várias técnicas de previsão a uma série temporal de arrecadação tributária. Além disso, estimula a utilização de técnicas de previsão qualitativas como forma de melhorar a acurácia dos resultados obtidos com os modelos estatísticos de previsão. O método de trabalho proposto apresenta uma seqüência estruturada de passos para a realização da previsão de receitas tributárias. Os modelos propostos foram testados com dados de arrecadação de ICMS no estado do Paraná.

[Barreto, 2005] propõe um método de classificação e seleção de contribuintes para fiscalização, fundamentado na previsão de seus comportamentos tributários por meio

de Modelos Lineares Generalizados Hierárquicos. A partir da estimação dos modelos propostos e da utilização de variáveis preditoras individuais de contribuintes e também de setores econômicos, foram devidamente quantificadas e explicadas as fontes de variabilidade nas respostas de interesse tributário enfocadas. Os modelos foram aplicados à classificação de contribuintes e avaliados, em termos de capacidade preditiva, pela estimação de suas medidas de validade interna.

Os trabalhos citados abordam modelos de comportamento para grupos de contribuintes. Entretanto, as administrações fazendárias normalmente tratam os contribuintes, em termos de fiscalização e predição de ações, de forma diferenciada, até porque a movimentação econômica tende a ser concentrada em um pequeno conjunto, grupo este que representa a maior parte da arrecadação de tributos, em grande parte reflexo da má distribuição de renda da própria população dos países subdesenvolvidos, dentre eles o Brasil.

3 MODELO COMPORTAMENTAL DE CONTRIBUINTES

Este capítulo apresenta um modelo computacional genérico de usuários, capaz de simular o comportamento do mesmo através da mineração dos seus dados de ações passadas. Em seguida, esse modelo é instanciado para os contribuintes de sistemas tributários, apresentando-se o micro-modelo e o macro-modelo comportamental dos mesmos.

3.1 Modelo genérico

Basicamente, modelar o usuário significa adquirir seu modelo, representar o seu conhecimento e atualizá-lo. A aquisição deste modelo pode ser feita de forma explícita (usuário informa como ele é), implícita (sistema verifica como usuário é) ou híbrida, combinando as duas formas anteriores [Kobsa, 1999][Sharma, 2001].

Não menos importante que a aquisição do conhecimento do usuário é a escolha da forma de representação desse conhecimento. Problemas distintos podem levar a formas de representação distintas, pois se deve escolher aquela que melhor se adapte ao problema em questão. Exemplos de formas de representação são redes *bayesianas*, representação baseada em lógica, parâmetros lineares, lógica *fuzzy*, redes neurais e ontologias [Kobsa, 1999]

Um modelo genérico de um usuário, baseado no modelo de agentes BDI e conforme proposto em [Silva, 2003], pode ser visualizado na figura a seguir.



Figura 7 - Modelo genérico de usuário

3.1.1 Estado Interno

Neste modelo, o estado interno é onde são armazenadas as propriedades dos usuários relacionadas com seus objetivos, motivações e situação atual, podendo ser interpretado como a mente do usuário. Estas propriedades são definidas em decorrência das características que se deseja incorporar aos usuários.

3.1.2 Percepção

Para que os usuários possam interagir com um ambiente de forma autônoma, decidindo qual ação adotar a cada momento, é necessário que eles sejam capazes de perceber o ambiente no qual estão inseridos. Uma ação somente pode ser executada pelo usuário após uma análise da situação ao seu redor.

3.1.3 Modelagem Comportamental

O mecanismo de ação autônoma dos usuários é implementado através de comportamentos. Os comportamentos definem como os usuários devem agir e/ou reagir para realizar determinadas tarefas no ambiente. Assim, a modelagem de comportamentos é uma das etapas mais importantes no desenvolvimento de um sistema de modelagem de usuários. A modelagem de comportamentos consiste na identificação e análise das ações que serão reproduzidas pelos usuários.

É importante compreender que a modelagem de comportamentos é dependente do fenômeno que está sendo simulado. Por exemplo, um modelo de comportamentos para a simulação da interação de contribuintes com o fisco, por exemplo, dificilmente será adequado para simular a interação de negociadores em uma bolsa de valores.

Apesar da modelagem do comportamento do usuário ser dependente do tipo de usuário a ser modelado, pode-se definir um conjunto de comportamentos genéricos que podem ser utilizados. Estes comportamentos genéricos definem ações básicas e podem ser associados para a criação de outros comportamentos mais complexos e específicos.

3.1.4 Modelo Computacional Proposto

Levando-se em consideração os dados que descrevem as interações do usuário com o ambiente ao longo do tempo, um modelo computacional de usuário pode ser representado como o conjunto:

$$M = \{Vei, Pp, Vei(t)\}. \quad (4)$$

Onde:

- Vei é um vetor de atributos que definem uma fotografia do usuário em um instante determinado (estado interno);
- Pp é um pacote de procedimentos de interação do usuário com o ambiente externo (percepção);
- Vei(t) corresponde a uma função que determina o vetor Vei em um instante “t”, obtida através de técnicas que permitam prever valores futuros com base em valores passados (comportamento).

3.2 Micro-Modelo de Contribuinte

O modelo básico de comportamento de contribuintes, segundo [Segura, 2006], propõe que a decisão de pagar ou não pagar um imposto resulta da comparação entre os custos e os benefícios.

Em sua apresentação mais simples, o benefício coincide com a quota devida caso o contribuinte deixe de pagar o imposto; e o custo está representado pelo valor do risco assumido neste caso, equivalente ao produto da quantidade que o contribuinte teria que pagar se fosse descoberto (a quota devida mais as sanções) multiplicado pela probabilidade de que isto ocorra. Ou seja:

- $CUSTO = (TRIBUTO\ DEVIDO + SANÇÃO) \times PROBABILIDADE\ DE\ SER\ DESCOBERTO$
- $BENEFÍCIO = TRIBUTO\ DEVIDO$

De acordo com este modelo, se o benefício for considerado superior ao custo, o contribuinte decidirá não pagar o imposto, isto é, decidirá fraudar, enquanto que se inclinará por pagar o imposto caso os custos a assumir resultem superiores ao benefício. Assim sendo:

- $SE\ BENEFÍCIO > CUSTO \Rightarrow FRAUDE$

Neste modelo básico, três variáveis fundamentais explicam a conduta dos contribuintes frente ao imposto: sua quantia (valor do imposto), a probabilidade de que a

administração tributária descubra o não cumprimento da obrigação e as sanções. A seguir serão tratadas tais variáveis.

3.2.1 Valor do imposto (V_i)

O valor do imposto constitui, conforme citado anteriormente, o benefício imediato que o contribuinte obteria se decidisse evadir o imposto. Em uma apresentação mais elaborada, caberia somar os custos de cumprimento que também poderiam ser evitados.

Nesta apresentação simplificada, uma elevação dos tipos impositivos, permanecendo iguais as demais coisas, iria aumentar em valores absolutos a diferença entre benefícios e custos e poderia estimular a fraude, tal como parecem confirmar os trabalhos empíricos a respeito [FRANZONI, 1998].

É notório observar que tal variável apresenta um relativo grau de dependência em relação a duas outras variáveis, Entradas (E) – que corresponde a todos os insumos acolhidos pelo contribuinte e Saídas (S) – que corresponde a tudo aquilo gerado e exteriorizado por tal contribuinte. Uma forma genérica de definir essa dependência é através da equação

$$V_i = (\sum S * p_s - \sum E * p_e) * p,$$

onde p_s corresponde às alíquotas de bases de cálculo para exportações, p_e às alíquotas de bases de cálculo para importações e p à alíquota para cálculo do valor imposto. Essa é uma forma genérica e simplificada de apresentar tal equação, sendo que os detalhes referentes a cada imposto podem ser verificados em [CTN, 1966].

3.2.2 Probabilidade de ser descoberto (PSD)

A segunda variável considerada é a probabilidade de ser descoberto (PSD). Esta variável pode ser associada com a probabilidade de ser auditado, medida pela relação existente entre o número de auditorias realizadas e o número de contribuintes, uma probabilidade que, normalmente, apresenta valores muito baixos, não superiores a 5%.

Com uma probabilidade média tão baixa de ser descoberto, poder-se-ia chegar à conclusão de que qualquer contribuinte racional deixasse de cumprir as obrigações. Entretanto, através de outros instrumentos que não a auditoria, pode-se aumentar essa probabilidade em muitos pontos percentuais, o que ocasiona um menor nível de omissão nas informações prestadas ao fisco.

Em qualquer administração tributária moderna, existem outros instrumentos de verificação das declarações que, embora não tenham a potência da auditoria, nem por isso deixam de ser relevantes. Como exemplo, tem-se os sistemas de retenções e pagamentos, os cruzamentos de informação e a utilização de “modelos de resultados” para selecionar declarações atípicas. Todos estes instrumentos permitem detectar fraudes e, portanto, devem aumentar o valor da PSD para bem mais que apenas 5%.

Em segundo lugar, a PSD pode ser um caso de informação assimétrica e, certamente, é freqüente que os contribuintes sobrevalorem extraordinariamente esta probabilidade. A impressão de que todo dado em posse da Administração Tributária, ou simplesmente a seu alcance, será efetivamente utilizado na verificação das declarações é bastante comum mesmo em países onde a Administração Tributária está longe de dispor de tal capacidade.

Em terceiro lugar, é muito evidente que a PSD não é a mesma para todos os contribuintes. A PSD para um empregado cujo salário está submetido à retenção é claramente muito superior a que corresponderia a um profissional liberal ou a um pequeno empresário. Apenas como exemplo, no caso dos Estados Unidos, estima-se que os assalariados submetidos à retenção chegam a declarar voluntariamente 99% de suas rendas, enquanto que no caso de alguns trabalhadores autônomos esta proporção pode cair até os 19%. Resulta portanto, com as hipóteses do modelo básico, que a opção por defraudar seja muito mais freqüente entre os profissionais liberais e pequenos empresários que no caso dos empregados, como acontece realmente.

3.2.3 Sanções (Vs)

A última variável considerada no modelo são as sanções. Em seu tratamento analítico, as sanções e a PSD aparecem como variáveis intercambiáveis. Formalmente, pois, a administração tributária poderia decidir gastar menos reduzindo a PSD, e compensá-la com um aumento no nível das sanções. Entretanto, de uma perspectiva aplicada, poucos aceitariam a substituição de ambas as variáveis além de margens estreitas. É que nem para a administração tributária nem para os cidadãos em geral, o volume arrecadado pode ser entendido como independente do padrão distributivo. Evidentemente, não é o mesmo arrecadar de todos os contribuintes, que contar com uma fraude ampla e um segmento de contribuintes pagando plenamente.

3.2.4 Modelo Comportamental

Levando-se em consideração a teoria do modelo básico comportamental de contribuintes, um modelo matemático que descreva a forma como se comporta um contribuinte ao longo do tempo pode ser dado pelo seguinte vetor:

$$v_{cc} = \{Ef, Ec, Sf, Sc, Vif, Vic, PSD, Vs\}.$$

O vetor v_{cc} representa o comportamento do contribuinte, sendo composto pelos seguintes elementos:

- Ec é o somatório das entradas declaradas pelo contribuinte ao fisco;
- Ef é o somatório das entradas do contribuinte percebidas pelo fisco;
- Sc é o valor das saídas declaradas pelo contribuinte ao fisco;
- Sf é o somatório das saídas do contribuinte percebidas pelo fisco;
- Vif é o imposto calculado sobre os valores de entrada e saída percebidos pelo fisco
- Vic é o imposto calculado sobre os valores de entrada e saída declarados pelo contribuinte
- PSD a probabilidade do contribuinte em ser descoberto no caso de omissão de informações corretas;
- Vs corresponde ao valor da sanção aplicada pelo fisco ao contribuinte, no caso de detecção de fraude.

Diferenças de valores entre os atributos Ec e Ef ou entre Sc e Sf indicam que existe fraude para tal contribuinte. A PSD é, dessa forma, diretamente proporcional às variações de valores ΔE e ΔS na maioria dos casos. Essa relação de proporcionalidade pode ser alterada em função de alguns critérios para cruzamento de dados dos contribuintes ou auditoria nos mesmos, tais como aleatoriedade na escolha dos contribuintes a serem

cruzados em função de carência de recursos computacionais. Nesses casos, a proporção direta pode deixar de existir. O valor da sanção aplicada corresponde ao somatório dos valores da variação $\Delta V_i = V_{if} - V_{ic}$ com a penalidade aplicada de acordo com legislação específica de cada imposto [CTN, 1966].

Ao se modelar o comportamento do contribuinte ao longo do tempo, pode-se obter o vetor V_{cc} em um instante t no futuro, ou seja, $V_{cc}(t)$. A utilização de técnicas de *forecasting*, ou seja, técnicas utilizadas quando da existência de um histórico de dados, mostra-se muito atraente neste caso, pois de posse do histórico dos valores de E_f e E_c , S_f e S_c , V_{if} e V_{ic} , PSD e V_s , pode-se inferir valores futuros de tais dados, obtendo em determinado instante o vetor comportamento do contribuinte - v_{cc} . Dessa forma, encontrar $v_{cc}(t)$ significa encontrar os valores $E_f(t)$, $E_c(t)$, $S_f(t)$, $S_c(t)$, $V_{if}(t)$, $V_{ic}(t)$, $PSD(t)$ e $V_s(t)$, isto é:

$$v_{cc}(t) = \{E_f(t), E_c(t), S_f(t), S_c(t), V_{if}(t), V_{ic}(t), PSD(t), V_s(t)\}$$

3.3 Macro-Modelo de contribuintes

O estudo de micro-modelos para compreensão de macro-modelos tem sido amplamente utilizado nos últimos anos na literatura [CLARAMUNT e JIANG]. A computação evolutiva [BÄCK, 1993][NOGUEIRA, 1999], por exemplo, tem mostrado que o estudo de micro-processos biológicos tem trazido grandes resultados no estudo de problemas mais amplos, como a resolução de equações; tem-se ainda o exemplo das redes neurais, onde através da simulação de pequenos processos ocorridos no interior de um cérebro artificial, consegue-se uma ampla gama de aplicações práticas, tais como previsão e reconhecimento de padrões [BISHOP, 1996].

De forma semelhante, através do conhecimento do comportamento de contribuintes individualmente, pode-se chegar a um comportamento global de uma determinada entidade fazendária. Isso pode ser obtido através do somatório dos comportamentos individuais. Dessa forma, o comportamento tributário da entidade pode ser dado pela equação:

$$v_{cs}(t) = \sum v_{cc}(t),$$

com $v_{cc}(t) = \{Ef(t), Ec(t), Sf(t), Sc(t), Vif(t), Vic(t), PSD(t), Vs(t)\}$ e

$$v_{cs}(t) = \{\sum Ef(t), \sum Ec(t), \sum Sf(t), \sum Sc(t), \sum Vif(t), \sum Vic(t), \sum PSD(t)/n, \sum Vs(t)\},$$

onde n é o número de contribuintes modelados.

Através do modelo global de comportamento, podem-se realizar simulações do tipo *what-if*, onde a variação de alíquotas para o cálculo dos valores Vic e Vif , por exemplo, podem ser analisados. O modelo contempla ainda uma forma de previsão de arrecadação, entradas e saídas, e nível de inadimplência por parte dos contribuintes.

O tamanho do grupo de contribuintes (n) pode ser determinado pela identificação do subgrupo e das características chaves dos contribuintes deste grupo, fazendo uma analogia com a representação do conhecimento dos usuários através de estereótipos (vide seção 2.1.1).

4 MODELO DE AGENTES PARA SOCIEDADE FAZENDÁRIA

A abordagem de Sistemas Multiagentes (SMA), introduzida pela Inteligência Artificial Distribuída (IAD), adicionou a idéia de Trabalho Cooperativo, onde um grupo de agentes (sociedade de agentes artificiais e humanos) trabalha em conjunto visando alcançar a um objetivo comum e onde cada um é responsável por realizar tarefas diferentes. Desta forma, alcança-se um alto grau de flexibilidade e adaptação a mudanças no ambiente e nos problemas a serem resolvidos.

Neste capítulo será descrita uma arquitetura para a sociedade de agentes que farão as previsões de comportamento para os contribuintes fazendários individualmente, bem como o projeto interno destes agentes.

4.1 ARQUITETURA

Para arquiteturas de sociedades de agentes, a utilização de ferramentas computacionais que forneçam um ambiente interativo e fácil de usar, onde o pesquisador possa ajustar o software para atender às suas necessidades, sem a necessidade de programação de código é algo bastante útil. De posse dessas ferramentas, os pesquisadores podem direcionar seus esforços para o estudo e criação de seus modelos, principalmente para a modelagem comportamental. Segundo Jiang (1999), o software StarLogo é um exemplo deste tipo de ferramenta.

Entretanto, não foram encontrados exemplos de ferramentas que operassem integradas com dados de Sistemas Tributários. Desta forma, para implementar o modelo em questão foi necessário desenvolver completamente as ferramentas (na forma de protótipos) que pudessem simular os modelos propostos neste trabalho. Para isso, foi

utilizada uma arquitetura simplificada para o desenvolvimento de ferramentas baseadas em agentes.

Diz-se que a arquitetura adotada é simplificada porque não considera algumas questões discutidas na literatura, tais como: comunicação entre agentes, mecanismos de memória, adaptabilidade ou aprendizado, etc. Entretanto, esta arquitetura pode ser facilmente expandida para atender a algumas destas questões sem grandes modificações.

A arquitetura de um sistema baseado em agentes define seus componentes e como eles devem estar relacionados, a fim de que os objetivos especificados para o modelo sejam atingidos. Como a arquitetura está intimamente relacionada com o conceito de agentes, foi natural adotar-se uma arquitetura baseada em 3 camadas. As camadas da arquitetura foram assim denominadas: Camada de Ambiente, Camada de Objetivos e Motivação e Camada de Ação. Elas formalizam como os componentes do modelo de agentes (estado interno, percepção e comportamentos) devem atuar durante a simulação. A Figura abaixo ilustra a arquitetura adotada.

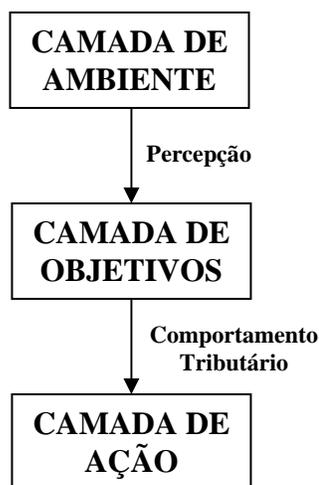


Figura 8 - Arquitetura simplificada utilizada para simulação

A Camada de Ambiente representa o ambiente de simulação propriamente dito englobando todos os agentes. Esta camada é consultada no início de cada ciclo da simulação, pelo módulo de percepção, com o objetivo de obter informações da região ao redor dos agentes, que serão utilizadas nas outras camadas para selecionar as ações que serão adotadas.

A Camada de Objetivos e Motivação implementa a “inteligência” dos agentes. A finalidade desta camada é analisar as informações obtidas da Camada de Ambiente e, baseado nas motivações e objetivos do estado interno, selecionar o comportamento mais apropriado para o momento. Nesta camada, adota-se o conceito de situação para qualquer arranjo de eventos percebidos no ambiente que implique na seleção de um determinado comportamento.

A Camada de Ação é responsável pela execução do comportamento selecionado na Camada de Objetivos e Motivação. A forma como os comportamentos serão executados depende das características individuais do agente (séries históricas dos atributos que definem o vetor comportamento). Com isso, dois agentes podem adotar um mesmo comportamento, mas apresentarem efeitos diferentes, devido às suas diferentes características.

A implementação da arquitetura foi bastante facilitada pela correlação entre suas camadas e a estrutura dos agentes (percepção, estado interno e comportamentos). A figura a seguir mostra a arquitetura adotada em uma perspectiva mais próxima da implementação.

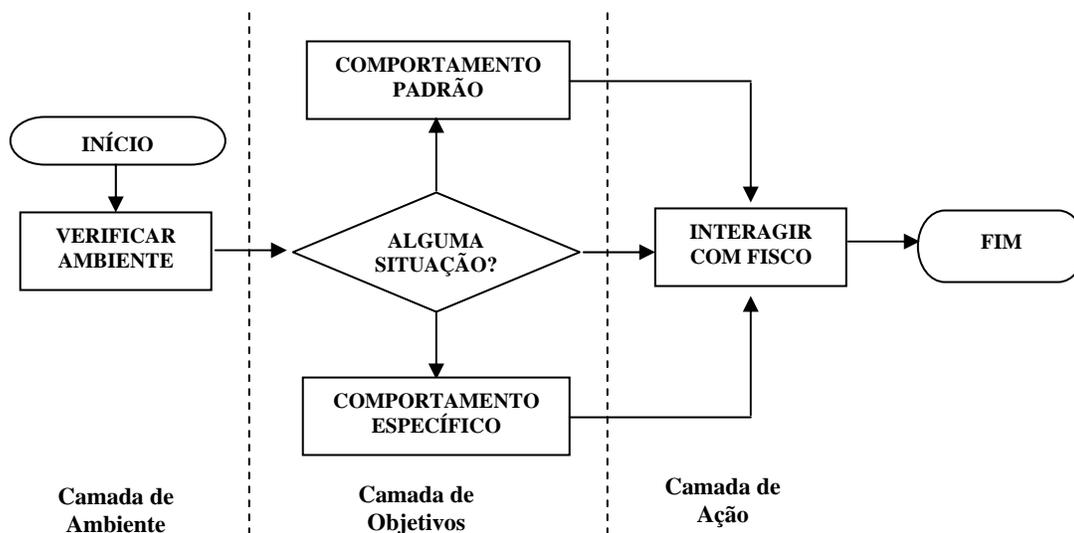


Figura 9 - Arquitetura detalhada utilizada para simulação

4.2 MODELAGEM DO AGENTE CONTRIBUINTE

Esta seção mostra o modelo de agente responsável por simular o comportamento de cada grupo de contribuintes. O termo grupo está sendo usado uma vez que um agente ora pode simular o comportamento de apenas um contribuinte, ora o comportamento de um grupo de contribuintes. No primeiro caso, o grupo é dito unitário. Estes grupos são relacionados ao nível de agregação detalhado na seção (2.3.1.2.2)

4.2.1 Conjunto modelo do contribuinte – M

Levando-se em consideração a modelo genérico de usuários abordado na seção anterior, o conjunto modelo M_c de um contribuinte é dado por

$$M_c = \{V_{eic}, P_{pc}, V_{eic}(t)\}. \quad (5)$$

$$P_{pc} = \{\text{“perceber PSD geral”}\}. \quad (6)$$

$$V_{eic}(t) = \{E_f(t), E_c(t), S_f(t), S_c(t), V_{if}(t), V_{ic}(t), PSD(t), V_s(t)\}. \quad (7)$$

A seguir serão comentados cada um dos elementos do conjunto M_c .

4.2.2 Vetor de estado interno – V_{ei}

O vetor de estado interno do agente em questão é o vetor $v_{cc} = \{E_f, E_c, S_f, S_c, V_{if}, V_{ic}, PSD, V_s\}$ descrito na seção 3.2.4.

Neste modelo, E_c é o somatório das entradas declaradas pelo contribuinte ao fisco e E_f é o somatório das entradas do contribuinte percebidas pelo fisco. S_c corresponde ao valor das saídas declaradas pelo contribuinte ao fisco e S_f é o somatório das saídas do contribuinte percebidas pelo fisco. V_{if} é o imposto calculado sobre os valores de entrada e saída percebidos pelo fisco. V_{ic} é o imposto calculado sobre os valores de entrada e saída declarados pelo contribuinte e PSD a probabilidade do contribuinte em ser descoberto no caso de declaração incorreta de informações. Por fim, V_s corresponde ao valor da sanção aplicada pelo fisco ao contribuinte, no caso de detecção de fraude.

4.2.3 Pacote de percepções - Ppc

A modelagem de percepções do agente contribuinte pode ser simplificada pela capacidade do contribuinte em adaptar sua variação de valor do imposto pago (ΔV_i) em relação à PSD de outros contribuintes. Apesar de não ter acesso aos dados exatos sobre os percentuais totais dos contribuintes auditados, através da comunicação informal, os contribuintes têm uma breve noção sobre este número. Assim, quando a PSD geral aumenta, o contribuinte tende a não omitir informações falsas ao fisco.

4.2.4 Funções de inferência – $V_{ei}(t)$

É importante observar que, uma vez que as inferências serão feitas para cada atributo do vetor e de forma individual por contribuinte, deve-se levar em consideração métodos de previsão que possam ser aplicados tanto a grandes massas de dados históricos (caso dos contribuintes que iniciaram suas atividades econômicas há mais tempo) como a pequenas e médias massas de dados históricos (contribuintes que iniciaram suas atividades econômicas mais recentemente). Isso devido à utilização da arquitetura em que cada contribuinte possui seu próprio modelo comportamental, modelo este obtido através de técnicas que permitem a previsão de cada atributo do vetor V_{ei} .

O modelo comportamental em si pode ser simplificado pela utilização de séries temporais para cada atributo pertencente ao vetor V_{ei} . Assim, um método de previsão de valores futuros baseados em séries históricas pode ser utilizado.

4.2.5 Comportamento do Agente

É importante considerar a forma de raciocínio descrita na seção 3.2. Assim, o agente que representa o grupo sendo modelado deve agir conforme o teste:

Se $Vif(t) > (Vs(t) + Vif(t)) \times PSD$ então

$$Vic(t) = \text{previsão}(Vic(t))$$

Senão

$$Vic(t) = Vif(t)$$

Neste caso, caso o produto da PSD pela soma dos valores da sanção e do valor do imposto devido sejam superiores ao valor do imposto devido (benefício é maior que custo), o agente pode fraudar, sendo o valor do imposto a ser pago definido pela aplicação de técnicas de previsão na base de dados histórica do atributo Vic. Entretanto, se o custo for superior ao benefício, o agente optará por não fraudar, pagando exatamente o valor devido (Vif).

Este raciocínio é estendido para os demais atributos do vetor Vei , com os valores informados pelo contribuinte assumindo os valores detectados pelo fisco, ou seja, $Xc = Xf$, onde X é qualquer um dos atributos do vetor Vei .

5 ESTUDO DE CASO: SEFAZ-MA

O modelo comportamental detalhado na seção anterior foi testado na Secretaria de Fazenda do Estado do Maranhão (SEFAZ-MA), com o objetivo de validar-se tal proposta.

A SEFAZ-MA possui como núcleo da sua estrutura de tomada de decisão um DataWarehouse implementado no SGBD Oracle. Neste banco de dados as séries necessárias para a implementação do modelo proposto no capítulo 4 apresentam-se na forma de tabelas relacionais. A seguir é mostrado o diagrama utilizado para a obtenção dos dados das séries históricas para cada contribuinte.

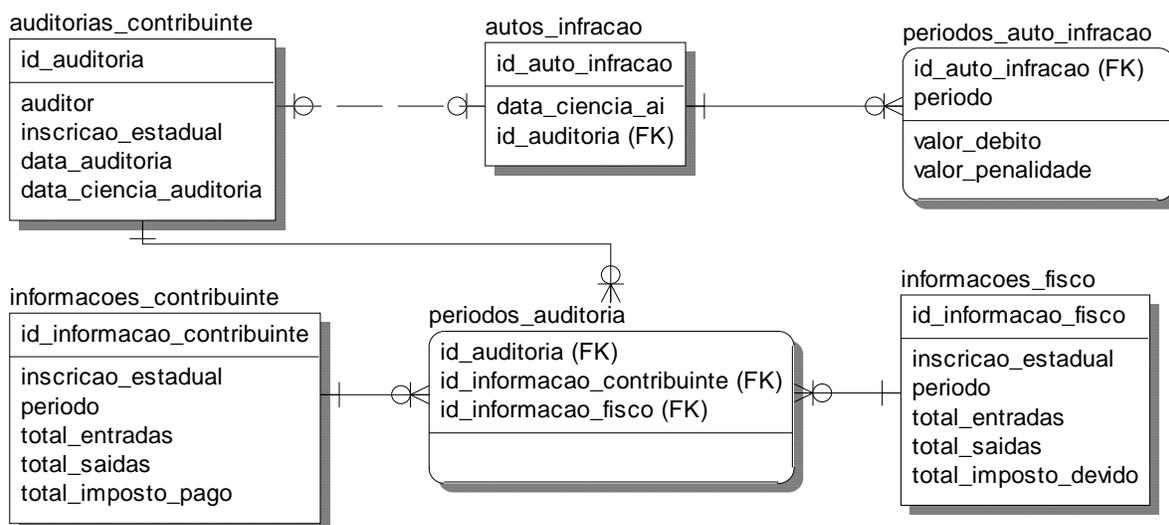


Figura 10 - Modelo relacional para extração das séries comportamentais

A tabela a seguir relaciona os atributos das tabelas do modelo anterior com os atributos do vetor comportamento descrito no capítulo 4.

Atributo do Modelo	Localização no Datawarehouse
Ef	informacoes_fisco.total_entradas
Ec	informacoes_contribuinte.total_entradas
Sf	informacoes_fisco.total_saídas
Sc	informacoes_contribuinte.total_saídas
Vif	informacoes_fisco.total_imposto_devido
Vic	informacoes_contribuinte.total_imposto_pago
PSD	Quantidade(período_auditoria)/ Quantidade(períodos_informações_contribuinte)
Vs	Soma(valor_penalidade)

Tabela 1 – Atributos do modelo versus datawarehouse

O tempo será avaliado neste caso como um número de mês, uma vez que o tributo modelado foi o ICMS (Imposto sobre Circulação de Mercadorias e Serviços de Transporte), sendo que os contribuintes são obrigados a declararem este imposto mensalmente [CTN, 1966]. O tempo t é então mapeado para a coluna período das tabelas do modelo em questão. Assim, a PSD(200412), por exemplo, leva em consideração a o quociente entre a quantidade de períodos de auditoria e a quantidade de períodos declarados para o contribuinte até o mês 200412. Esta, que pertence ao vetor de atributos que representam o estado interno do contribuinte, não deve ser confundida com a PSD geral, que representa a percepção do contribuinte sobre como fisco vem agindo em relação a fraudes. No caso desta última, a PSD(200412) corresponde ao quociente entre a quantidade de períodos de auditoria e a quantidade de períodos declarados para todos contribuintes até o mês 200412.

Para a simulação computacional dos modelos estudados, não foi encontrada ferramenta que pudesse ser utilizada de forma integrada com sistemas tributários. Desta forma, foi necessário que se codificasse um sistema que implementasse o modelo em

questão. A figura a seguir mostra a arquitetura de implementação do sistema para o caso da SEFAZ-MA.

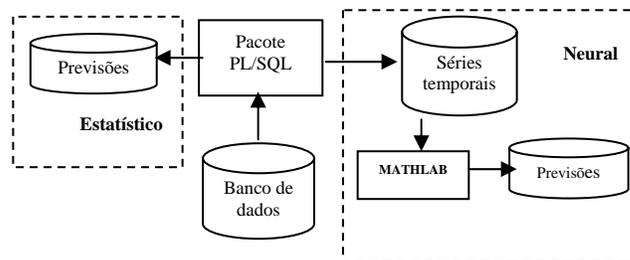


Figura 11 - Arquitetura de implementação do sistema de previsão comportamental

O núcleo do sistema, responsável por coletar os dados das séries temporais no banco de dados da SEFAZ foi implementado utilizando-se a ferramenta PL/SQL, proprietária da Oracle Corporation, uma vez que a SEFAZ-MA possui ambientes Operacional e de DataWarehouse que usam este SGBD, com os dados das séries históricas já encontrados nos formatos desejados, além do excepcional desempenho do PL/SQL em ambientes Oracle. A seguir será detalhada a forma de implementação de cada um desses componentes do sistema.

5.1 IMPLEMENTAÇÃO

A parte de inferência do modelo comportamental dos contribuintes foi implementada de duas formas distintas: usando o modelo estatístico de Holt e usando redes neurais artificiais (RNA's). Para o modelo estatístico, as previsões foram geradas diretamente pelo módulo PL/SQL. Já no modelo neural, os dados das séries históricas foram exportados para arquivos texto pelo módulo PL/SQL e lidos pela ferramenta MATHLAB para a geração das previsões.

5.1.1 Módulo estatístico

Para a implementação do módulo estatístico em questão considerou-se utilizar um modelo que pudesse ser aplicado tanto a grandes quanto a pequenas massas de dados. Isto porque na população de contribuintes a serem modelados há uma variação na quantidade de dados disponíveis para cada contribuinte. Os modelos de Box-Jenkins, conforme tratado no capítulo 2, necessitam de uma amostra com no mínimo 50 observações. Um modelo simples, fácil de ajuste e que apresenta boa precisão nos resultados é o Modelo de Holt (também mostrado no capítulo 2). O método estatístico implementado considerou este último modelo para a previsão dos comportamentos individuais dos contribuintes, implementado as equações detalhadas no capítulo 2 em PL/SQL.

5.1.2 Módulo neural

De forma a modelar o problema proposto, foi utilizada uma rede neural com três camadas: entrada, oculta e saída, conforme detalhado no capítulo 2. Para tanto, foram utilizadas duas abordagens.

Na primeira abordagem, usando a topologia de [Varfis e Versinom, 1990] discutida no capítulo 2, foi feita a previsão de cada atributo do vetor $Veic(t) = \{Ef(t), Ec(t), Sf(t), Sc(t), Vif(t), Vic(t), PSD(t), Vs(t)\}$ de forma individual. Para cada um destes atributos utilizou-se como variáveis de entrada: os onze últimos valores da série (dos instantes $t-1$ a $t-11$) como unidades atrasadas; os valores nos instantes $t-12$, $t-24$ e $t-36$ como unidades anuais e mais doze unidades binárias para captar a sazonalidade da série. Esses valores foram 10000000000 para janeiro, 010000000000 para fevereiro, até 000000000001 para dezembro.

Na camada oculta o número de neurônios variou de 14 a 53, considerando o comentário feito no capítulo 2. A função escolhida foi a sigmoïdal (logística), por sua natureza não linear, assim como também pelo fato de sua saída ser formada apenas por valores positivos, tendo em vista que os valores dos atributos do modelo comportamental também o são, sem exceção.

A camada de saída foi implementada com apenas um neurônio, já que na saída a rede fornecerá apenas um valor, que é a previsão para o atributo no mês desejado. A função de transferência desta camada foi a linear.

A figura a seguir ilustra a rede neural usada nessa primeira abordagem. Nela, está mostrada a rede neural para a previsão de valores do atributo $Vic(t)$ (série com os valores de imposto pagos pelo contribuinte). Para os demais atributos do vetor $Vei(t)$, foi utilizado o mesmo procedimento.

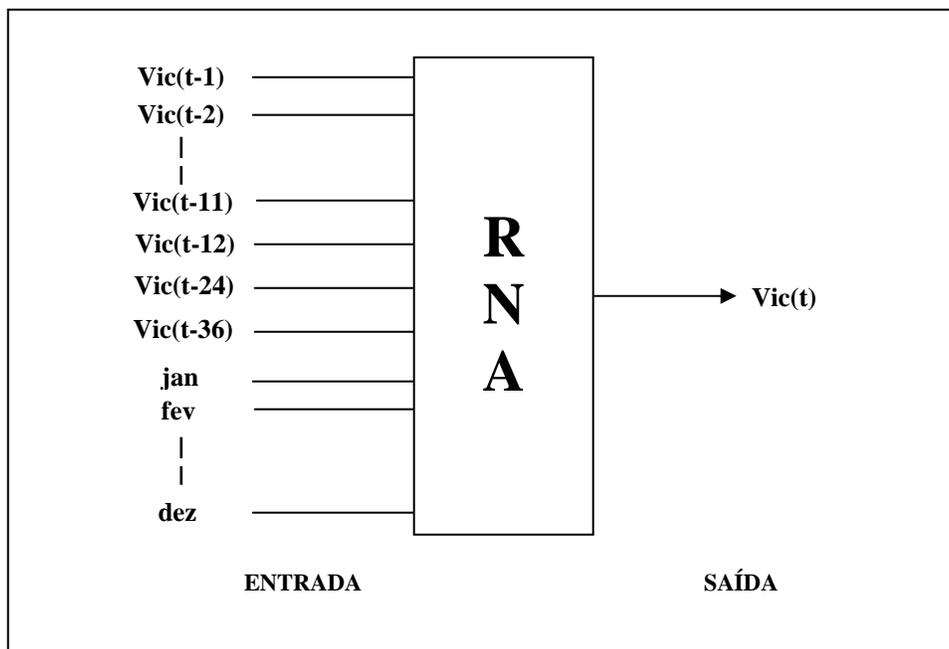


Figura 12 - RNA utilizada no módulo neural de previsão

A segunda abordagem, em vez de prever individualmente cada atributo, fez a previsão do vetor Vei para todos os atributos simultaneamente. A camada de entrada recebeu então os valores das variáveis de $t-1$ a $t-11$ e os valores $t-12$, $t-24$ e $t-36$ para cada atributo do vetor Vei . Assim a camada de entrada foi alimentada com 8×14 entradas dos atributos mais as 12 entradas binárias para captação da sazonalidade, resultando num total de 124 variáveis de entrada. O número de neurônios da camada intermediária variou de 66 a 249. A camada de saída foi implementada com 8 neurônios, já que na saída a rede fornecerá agora não mais um, mas 8 valores de previsão (Ef , Ec , Sf , Sc , Vif , Vic , PSD e Vs). Para a padronização dos dados, foi utilizado o escore z , que representa o número de desvios padrão que um valor dista da média do conjunto de valores.

Foram apresentados à rede, para treinamento, dados sobre as variáveis de entrada e saída correspondentes ao período de janeiro de 2000 até dezembro de 2004 (matriz de entrada e matriz-alvo). Como teste/validação foram realizadas as previsões para os meses

de janeiro até dezembro de 2005, ou seja, três meses à frente. A seguir, os resultados obtidos foram comparados com os valores reais ocorridos naquele período e calculados os erros da previsão. É fundamental ressaltar que há sempre erro na estimação por simulação.

O modelo utilizado neste trabalho foi implementado em um servidor HP Proliant G3, com 4 processadores Intel Pentium Xeon 2.2Ghz, 8GB de memória RAM, sistema operacional linux RedHat AS 3.0, SGBD Oracle release 9.2.0.5. Para o módulo de redes neurais, foi utilizada a engine do software MATLAB versão 7.3.

5.2 RESULTADOS OBTIDOS

Foram feitas simulações do comportamento dos contribuintes do ICMS do Estado do Maranhão, tomando-se como medida de comparação para os métodos os valores globais de entradas, saídas, arrecadação e sanções aplicadas aos contribuintes por não cumprimento da obrigação tributária. Estes valores globais foram comparados ao somatório dos valores individuais previstos para cada contribuinte. Foram feitas simulações para previsão dos valores dos atributos do vetor V_{ei} dos contribuintes nos instantes $t+1$ e $t+6$, ou seja, 1 mês e 6 meses após o mês atual.

A seguir será feita uma comparação dos resultados obtidos com os módulos neural e estatístico para o atributo V_{ic} (série de valores de imposto pago pelo contribuinte) do vetor V_{ei} (vetor de estado interno). Em seguida, compara esses resultados com as abordagens tradicionais de modelos comportamentais que envolvem grupos de contribuintes, em vez de modelos comportamentais individuais.

No caso dos modelos comportamentais para grupos não unitários de contribuintes, foram feitas previsões usando três tipos de grupos: setor econômico, grupo de atividade econômica e faixa de faturamento, além do supergrupo composto por todos os contribuintes. Foram utilizados os dados dos 1000 contribuintes com maiores valores de arrecadação em 2004 e 2005 para os testes, o que corresponde a um percentual de 94% da arrecadação do estado.

A lista completa de resultados obtidos para todos os atributos do vetor *Vei* encontra-se nos anexos.

5.2.1 Módulo estatístico versus Módulo neural

A tabela a seguir mostra os valores dos erros médios obtidos para as previsões das séries históricas para o ano de 2005.

Mês	Holt	RNA	Holt	RNA
	1 passo à frente		6 passos à frente	
Jan	4,98	6,72	27,55	25,84
Fev	4,64	9,85	30,16	26,72
Mar	6,97	8,97	30,03	24,56
Abr	7,17	5,72	28,19	24,93
Mai	3,78	3,67	26,32	25,45
Jun	5,88	4,82	32,29	31,16
Jul	7,80	3,56	28,73	26,45
Ago	9,30	3,91	27,01	22,32
Set	7,73	3,51	26,98	22,74
Out	4,14	4,75	30,41	29,86
Nov	3,96	5,03	26,23	25,51
Dez	4,63	4,84	28,31	25,29
Média	5,92	5,46	28,52	25,90

Tabela 2 - Erro médio de *Vic* obtido (Holt versus RNA)

Para a rede neural, nos casos em que não houve movimento nos meses devido à inexistência de obrigação tributária naquela época (por exemplo, o contribuinte ainda não existia), os pesos sinápticos da rede neural foram fixados em zero.

Os erros médios obtidos apresentaram um percentual variando de 3,78% a 9,30% para o modelo de Holt um passo a frente, de 3,51% a 9,85% para a RNA um passo a frente, de 26,23% a 32,29% para o modelo de Holt 6 passos à frente e de 22,32% a 31,16% para a RNA 6 passos à frente.

Os resultados obtidos mostraram que para previsões de 1 passo a frente, ou seja, para o próximo mês, a RNA foi mais precisa, conforme verificado na tabela 2 (5,46% da RNA para 5,92% do modelo de Holt), mas o modelo de Holt foi superior à RNA em 6 dos 12 meses. Já nas previsões para 6 passos a frente, a RNA mostrou-se amplamente superior, mostrando-se neste caso mais precisa que o outro modelo em todos os meses previstos.

Os melhores resultados para a previsão foram os obtidos com a MLP treinada com o algoritmo *backpropagation* usando a primeira abordagem de rede neural discutida anteriormente e com os seguintes parâmetros:

Item	Solução
Função treino	TRAININGDM
Função aprendizagem	LEARNINGDM
Neurônios da camada oculta	17
Épocas de treino	2000
Limite de erro (goal)	0,001
Taxa de aprendizagem	0,01

Tabela 3 - Parâmetros da RNA utilizada

O número de épocas de treinamento e o limite para o erro (goal) representam os critérios de parada utilizados no treinamento.

5.2.2 Modelagem de grupos versus modelagem individual de contribuintes

Na tabela abaixo são mostrados os resultados obtidos utilizando-se previsão com os modelos de grupos de contribuintes.

Mês	SG	SE	GE	AE	FF	MI	SG	SE	GE	AE	FF	MI
	1 passo a frente						6 passo a frente					
Jan	6,89	6,15	5,99	5,19	5,65	4,98	33,12	31,06	30,70	30,82	31,49	27,55
Fev	5,74	5,81	4,71	4,98	4,54	4,64	31,40	34,40	34,85	30,30	33,92	30,16
Mar	6,85	6,92	5,94	7,08	4,93	6,97	32,39	33,71	33,23	29,77	33,92	30,03
Abr	5,35	7,80	6,19	6,37	6,51	7,17	34,30	30,58	30,99	29,00	31,05	28,19
Mai	7,17	6,53	8,09	5,88	5,94	3,78	32,67	30,82	31,14	31,35	30,94	26,32
Jun	5,88	5,57	5,59	3,76	6,38	5,88	31,02	29,75	29,86	26,25	29,66	32,29
Jul	7,80	5,33	6,48	7,04	7,29	7,80	32,16	33,22	32,89	26,75	33,42	28,73
Ago	8,86	5,15	5,65	7,28	8,60	9,30	31,57	30,43	29,69	29,69	30,23	27,01
Set	7,38	5,48	7,57	6,18	5,80	7,73	32,71	33,26	34,12	29,47	32,69	26,98
Out	4,28	5,82	6,53	6,85	6,55	4,14	32,01	34,27	33,65	27,04	34,84	30,41
Nov	5,52	7,21	5,86	5,50	5,99	3,96	33,38	31,19	32,39	27,65	31,12	26,23
Dez	4,63	6,98	4,93	5,31	6,37	4,63	30,93	32,87	31,88	28,13	33,03	28,31
Média	6,36	6,23	6,13	5,95	6,21	5,92	32,31	32,13	32,12	28,85	32,19	28,52

Tabela 4 - Erro médio de Vic obtido com modelo de Holt (grupal versus individual)

Mês	SG	SE	GE	AE	FF	MI	SG	SE	GE	AE	FF	MI
	1 passo à frente						6 passos à frente					
Jan	7,28	7,60	6,18	6,84	6,95	6,72	25,90	25,82	25,82	25,67	25,60	25,84
Fev	7,20	6,50	6,19	6,52	6,79	9,85	27,91	26,27	25,96	24,86	27,42	26,72
Mar	6,08	6,62	6,66	6,63	6,63	8,97	24,40	25,11	25,35	25,57	25,27	24,56
Abr	6,38	6,56	5,92	6,51	6,65	5,72	25,55	24,8	24,90	24,07	24,61	24,93
Mai	6,88	6,91	6,73	4,19	6,92	3,67	25,35	25,12	24,79	25,07	25,83	25,45
Jun	6,01	5,92	5,59	4,94	5,72	4,82	32,17	31,29	31,46	31,43	31,86	31,16
Jul	6,77	6,26	5,88	5,54	6,22	3,56	26,23	26,67	27,13	27,23	26,76	26,45
Ago	5,26	6,08	6,1	5,99	6,08	3,91	22,75	22,52	22,40	22,06	22,54	22,32
Set	5,24	5,07	4,53	4,72	4,50	3,51	23,43	24,79	24,25	24,32	23,60	22,74
Out	6,28	5,72	6,2	5,06	6,28	4,75	30,32	29,25	29,97	30,64	30,32	29,86
Nov	4,10	4,42	4,43	4,68	4,25	5,03	25,94	25,38	24,82	24,36	25,38	25,51
Dez	6,23	5,99	5,81	5,1	6,19	4,84	25,28	25,62	25,64	25,81	25,90	25,29
Média	6,14	6,13	5,85	5,56	6,10	5,45	26,27	26,05	26,04	25,92	26,26	25,90

Tabela 5 - Erro médio de Vic obtido com modelo de RNA (grupal versus individual)

Nas tabelas acima, SG significa o supergrupo de contribuintes (todos os 1000 contribuintes). SE representa a divisão dos contribuintes por setor econômico (primário, secundário e terciário), GE significa grupo econômico (atacadista, varejista, indústria, telecomunicações, energia elétrica), AE é a divisão dos contribuintes por atividade econômica (CNAE – Código nacional de atividade econômica), FF representa a divisão dos grupos por faixa de faturamento (foram utilizadas 50 faixas de faturamento). Finalmente MI significa modelo individual, ou seja, foi gerado um modelo de comportamento para

cada contribuinte, sendo os mesmos já mostrados na comparação do módulo estatístico com o módulo neural.

Avaliando-se os dados em questão percebe-se novamente que, nas previsões 1 passo a frente, a qualidade dos dados obtidos usando o modelo de Holt aproximou-se das redes neurais em todos os conjuntos de contribuintes. Nas previsões 6 passos a frente, mais uma vez a rede neural mostrou-se amplamente superior ao modelo de Holt, superando-a em todos os meses.

O fato mais importante é que é percebida uma tendência na melhoria da qualidade das previsões à proporção que se diminui o tamanho do conjunto de contribuintes em questão. Como pôde ser percebido, o melhor resultado foi obtido com o grupo unitário de contribuintes, ou seja, modelando individualmente seus comportamentos, para em seguida compor o comportamento geral da sociedade de contribuintes. Isso pôde ser observado tanto na modelagem de grupos usando os modelos de Holt, quanto na modelagem usando redes neurais.

6 CONCLUSÕES

Neste trabalho foi proposto um modelo comportamental genérico de usuários baseado em inferência de ações futuras a partir de ações passadas desses usuários. O modelo em questão foi aplicado a um tipo especial de usuários, os contribuintes de sistemas fiscais. O resultado desta proposta é um modelo comportamental baseado em funções de inferência comportamental. Foram propostas duas abordagens de obtenção dessas funções: uma baseada em métodos estatísticos clássicos e outra baseada em redes neurais. Os resultados desta proposta foram relatados aqui, aplicando-se o modelo comportamental de contribuintes proposto à Secretaria de Fazenda do Estado do Maranhão do Brasil.

É importante frisar que, apesar do objetivo deste trabalho não ter sido desenvolver um modelo de previsão de variáveis tributárias, os resultados obtidos mostraram que os erros percentuais médios obtidos, principalmente referentes à série arrecadação de ICMS, mostraram-se bastante aceitáveis, levando a indagações sobre um possível estudo para desenvolvimento de modelos de previsão global de arrecadação a partir das suas divisões em micro-modelos de previsão individuais.

Outro ponto a ser destacado é que o modelo adotado permite que o método de previsão utilizado seja facilmente trocado, dependendo da necessidade de considerações a serem feitas, assim como o nível de agregação dos dados a serem simulados. Conforme foi mostrado no trabalho, várias opções podem ser usadas como, em vez de modelar um agente por contribuinte, um agente por atividade econômica ou setor econômico, ou mesmo um agente para todos os contribuintes, assemelhando-se dessa forma aos modelos de previsão clássicos já utilizados na literatura.

Uma das principais dificuldades encontradas foi a reduzida quantidade de literatura sobre modelos formais de contribuintes fiscais. Outra dificuldade encontrada foi o desempenho do sistema devido a grande quantidade de dados analisada, além da comunicação entre o SGBD Oracle (extrator das séries) e o software MATHLAB (implementação das redes neurais).

Sugere-se como trabalho futuro trazer a rede neural para dentro do banco de dados com as séries históricas a serem usadas como entrada, melhorando assim o desempenho do modelo para o módulo neural, e resolvendo uma das principais dificuldades encontradas que foi a comunicação Oracle/MATHLAB. Outra sugestão seria o desenvolvimento de uma interface gráfica para a visualização da população de contribuintes interagindo com o fisco, bem como para a parametrização do modelo.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABELÉM, A. J. G.: Redes neurais artificiais na previsão de séries temporais. Dissertação de Mestrado - Pontifícia Universidade Católica, Rio de Janeiro, Brasil, (1994).
- APPELT, D. E., and POLLACK, M. E.: Weighted Abduction for Plan Ascription. *User Modeling and User-Adapted Interaction*, 2:1-25, 1992.
- ARMSTRONG, J. S.: Principles of forecasting: a handbook for researchers and practitioners, Kluwer, Philadelphia, (1999).
- ARRAES, Ronaldo de Albuquerque; CHUMVICHITRA, Pichai. Modelos Autoregressivos e Poder de Previsão: Uma Aplicação com o ICMS, Texto para discussão n 152, Universidade Federal do Ceará, Programa de Pós-Graduação em Economia, 1996.
- ARRAGON, Paul Van: Nested Default reasoning for User Modelling. PhD thesis, University of Waterloo, (1990).
- BÄCK, T. and SCHWEFEL, H.-P.. An overview of evolutionary algorithms for parameter optimization. *Evolutionary Computation*, 1(1):1-23, 1993.
- BALLIM, Afzal and WILKS, Yorick: Beliefs, stereotypes and dynamic agent modelling. *User Modelling and User-Adapted Interaction*, 1(1):33-65, (1991).
- BARRETO, Alexandre Serra: Previsão De Comportamento e Classificação de Contribuintes Tributários: Uma Abordagem por Modelos Lineares Generalizados Hierárquicos, Qualificação de Doutorado, Universidade Federal de Santa Catarina, Florianópolis-SC, Brasil, Setembro (2005).
- BARTO, A.G., *Cellular Automata as Models of Natural Systems*. The University of Michigan, 1975.
- BERNARDO, Jordânia Rosa. Análise da Arrecadação do ICMS do Estado de Roraima: Evolução e perspectiva de potencial. Universidade Federal do Rio Grande do Sul, Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Economia, 2001.
- BISHOP, C. M., *Neural Networks for Pattern Recognition*.. Oxford University Press, 1996.
- BOUNABAT, B. "A New Formal Approach for the Specification and the Verification of Multi-agent Reactive Systems Operating Modes". In: *Information Processing and Technology*, pp. 25-47, Nova Science Publishers, Inc. New York, EUA, 2001.
- BRAJNIK, G.; GUIDA G.; TASSO, C.. User modeling in intelligent information retrieval. *Information Processing and Management*, 23:305-320, 1987.
- BRAZ, Eugênio Rubens Cardoso: Um Modelo para Gerenciamento, Avaliação e Planejamento da Arrecadação de Tributos Estaduais, Qualificação de Doutorado, Universidade Federal de Santa Catarina, Florianópolis-SC, Brasil, Abril (2001).
- CAUDILL, M. (1990), *Neural Network Primer*, Miller Freeman Publications.
- CHIN, David Ngi: Intelligent Agents as a Basis for Natural Language Interfaces. PhD thesis, Computer Science Division (EECS), University of California, Berkeley, CA94720, jan, 1988.

- CLARAMUNT C., JIANG B., “A Qualitative Model for the Simulation of Traffic Behaviours in a Multi-lane Environment”, *Journal of Geographical Sciences*, v. 11, Special issue on geo-visualization, pp. 29-42, 2001.
- COCCARO, Suzana Menna Barreto. A Arrecadação do ICMS-RS: Um Enfoque Econométrico. Universidade Federal do Rio Grande do Sul, Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Economia, 2000.
- CORRÊA, W. R.; PORTUGAL, M. S. Previsão de séries de tempo na presença de mudança estrutural: redes neurais artificiais e modelos estruturais. *Economia Aplicada*, v.2, n.3, p.487-514, 1998.
- CORVALÃO, Eder Daniel: Previsão da arrecadação do imposto sob circulação de mercadorias e serviços em Santa Catarina: Aplicação da Abordagem Geral para Específico em Modelos Dinâmicos, Qualificação de Mestrado, Universidade Federal de Santa Catarina, Florianópolis-SC, Brasil, Setembro (2002).
- CSINGER, Andrew and POOLE, David: From utterance to belief via presupposition: Default reasoning in user-modelling. In *Proceedings of the Conference for Knowledge Based Computing Systems, KBCS-89*, pages 408–419, Bombay, India, December, (1989).
- CTN, 1966: Código Tributário Nacional, Secretaria da Receita Federal do Brasil, www.receita.fazenda.gov.br/Legislacao/CodTributNaci/ctn.htm, Brasília – DF, Brasil, Outubro, (1966).
- DIAS, G. P. P. Proposta de processo de previsão de vendas para bens de consumo. *Anais do XIX ENEGEP – CD-ROM*, Rio de Janeiro, 1999.
- ELLER, R. and CARBERRY, S.: A Meta-Rule Approach Modeling and User-Adapted Interaction, 2:7-53, 1992.
- FERREIRA, Roberto Tatiwa. Modelo de análise de séries temporais para previsã do ICMS mensal do Ceará. Universidade Federal do Ceará, Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Economia, 1996.
- FININ, T. W. and Drager, D.: A general user modeling system. *Proc. of the 6th Canadian Conference on Artificial Intelligence*, Montreal, Canada, pp. 24-29 (1986).
- FRANZONI, Luigi Alberto: Tax evasion and tax compliance, *Encyclopaedia of Law and Economics*, University of Bologna, Itália, September, (1998).
- FREIMAN, José Paulo e PAMPLONA, Edson de O.: Redes Neurais Artificiais na Previsão do Valor de Commodity do Agronegócio. V *Encuentro Internacional de Finanzas*. Santiago, Chile, 19 a 21 de janeiro de (2005).
- HILL, T. (1994). *Manufacturing strategy: text and cases*, 2^a ed., Irwin, Boston, MA.
- JENNINGS, A. and HIGUCHI, H. (1993). A User Model Neural Network for a Personal News Service. *User Modeling and User-Adapted Interaction* 3:1-25.
- JIANG B., 1999, “SimPed: Simulating Pedestrian Flows in a Virtual Urban Environment”, *Journal of Geographic Information and Decision Analysis*, v. 3, n. 1, pp. 21-30.

- JONES, Marlene and POOLE, David: An expert system for educational diagnosis based on default logic. In *Proceedings Expert Systems and their Applications*, pages 673–683, (1985).
- KOBSA, A.: A taxonomy of beliefs and goals for user models in dialog systems.-In: A. Kobsa and W. Wahlster (eds.), *User Models in Dialog Systems*. Springer-Verlag, Berlin, Heidelberg, pp. 52-68 (1989).
- KOBSA, A.: “Generic User Modeling Systems.” *User Modeling and User-Adapted Interaction* 11(1-2), 49-63 (2001).
- KOBSA, Alfred: *Personalised Hypermedia Presentation Techniques for Improving Online Customer Relationships*, GMD Report 66, (1999).
- KOBSA, A.: *User Modeling: Recent Work, Prospects and Hazards*. In: M. Schneider-Hufschmidt, T. Kühme, U. Malinowski, eds.: *Adaptive User Interfaces: Principles and Practise*. Amsterdam: North Holland Elsevier (1993).
- LIEBEL, Marlon Jorge: *Previsão de Receitas Tributárias – O Caso do ICMS no Estado do Paraná, Qualificação de Mestrado, Universidade Federal do Rio Grande do Sul - RS, Brasil, (2004)*.
- MAKRIDAKIS, S., WHEELWRIGHT, S. C. & HYNDMAN, R. J. *Forecasting – methods and applications*, 3^a ed., John Wiley, New York, 1998.
- MAYFIELD, J.: *Controlling Inference in Plan Recognition*. *User Modeling and User-Adapted Interaction*, 2:83-115, 1992.
- MENTZER, J. T. & COX Jr., J. E. *Familiarity, application, and performance of sales forecasting techniques*. *Journal of Forecasting*, v. 3, n. 1, p. 27-37, jan. 1997.
- MONTGOMERY, D. C., JOHNSON, L. A. & GARDINER, J. S. *Forecasting and time series analysis*, 2^a ed., McGraw-Hill, Inc., New York, 1990.
- NOGUEIRA, M.L.; SAAVEDRA, O. R. “Estratégias Evolutivas Aplicadas à Resolução de Otimização Multimodal”, *Simpósio Brasileiro de Automação Inteligente*, 1999.
- PERRAULT, C. R., Allen, J. F. and COHEN, P. R.: *Speech acts as a basis for understanding dialogue conference*. Report 78-5, Department of Computer Science, University of Toronto, Canada (1978).
- RASKUTTI, B., and ZUKERMAN, I.: *Generation and Selection of Likely Interpretations During Plan Recognition in Task-Oriented Consultation Systems*. *User Modeling and User-Adapted Interaction*, 1:323-353, 1991.
- RICH, Elaine: *Users are individuals: Individualizing user models*. *International Journal of Man-Machine Studies*, 18:199–214, (1983).
- RICH, Elaine: *User modeling via stereotypes*. *Cognitive Science*, 3:329–354, 1979.
- RUMELHART, David E.; HINTON, G. E.; WILLIAMS, R. J.: *Learning Internal Representations by Error Propagation*. In *Parallel Distributed Processing*, Cambridge : M. I. T. Press, v. 1, p. 318-362. (1986)

- SANDERS, N. R. & MANRODT, K. B. Forecasting practices in US corporations: survey results. *Interfaces*, v.24, n. 2, p. 92-101, mar. 1994.
- SARNER, Margaret H. and CARBERRY, Sandra: Tailoring definitions using a multifaceted user model. In *Proceedings of the Eighth Biennial Conference of the Canadian Society for Computational Studies of Intelligence*, pages106–113, (1990).
- SEGURA, José V. Sevilla: A arrecadação potencial como meta da administração tributária, Centro Interamericano de Administrações Tributárias – CIAT, 40a. Assembléia Geral, Florianópolis – SC, Brasil, (2006).
- SHAFER, G. (1976): *A Mathematical theory of evidence*, Princeton University Press, Princeton, NJ.
- SHARMA, Amit: *A Generic Architecture for User Modeling Systems and Adaptive Web Services*, In: *Workshop on E-Business & the Intelligent Web*. (2001).
- SILVA, Cláudio Antônio: *Modelagem comportamental para agentes autônomos em ambiente reais*, Qualificação de Mestrado, Universidade do Estado do Rio de Janeiro, Rio de Janeiro - RJ, Brasil, Setembro, (2003)
- SILVEIRA, Altair Lizete Queiroz da. *Um modelo de Previsão da Arrecadação do ICMS em Goiás*. Fundação Getúlio Vargas, Escola Brasileira de Administração Pública, Monografia submetida ao curso de Administração Pública, Goiânia, 2000.
- TOMPKINS, J. A., BOZER, Y. A., FRAZELLE, E., et al. (1996). *Facilities planning*, 2^a ed., John Wiley, New York.
- TURBAN, E. *Decision support and expert systems: management support systems*. New York: MacMillan, 1993.
- VARFIS, A. , VERSINO, C., *Univariate Economic Time Series Forecasting*, Cambridge University Press, Cambridge, 1990.
- WOOLDRIDGE, M. *An Introduction to Multi-Agent Systems*, John Wiley & Sons Ltda. Chichester, Inglaterra, 2002.
- WOOLDRIDGE, M., and JENNINGS, N. *Agent Theories, Architectures and Languages: A Survey*, In *Proceedings of Workshop on Agent Theories Architectures, and Languages (ECAI 1994)*, *Lecture Notes in Artificial Intelligence*, AmsterSpringer-Verlag, pp. 01-32, 1994.

ANEXOS

ANEXO I – INTERFACE DO PACOTE COM CÓDIGO FONTE

```

create or replace package pck_comportamento_contribuinte is

    C_METODO_ESTADISTICO_01 constant number:=10; -- Modelo de Holt
    C_METODO_NEURAL_01     constant number:=20; -- Rede neural c/ saída
    simples
    C_METODO_NEURAL_02     constant number:=21; -- Rede neural c/ saída
    completa

    C_SG      constant number:=1; -- Modelagem com apenas um grupo
    C_SE      constant number:=2; -- Modelagem por setor econômico
    C_GE      constant number:=3; -- Modelagem por grupo econômico
    C_AE      constant number:=4; -- Modelagem por atividade econômica
    C_FF      constant number:=5; -- Modelagem por faixa de faturamento
    C_MI      constant number:=6; -- Modelagem por contribuinte
    (individual)

    -- Vetor de estado interno do grupo
    -- O grupo pode ser um dos tipos C_SG, C_SE, C_GE, C_AE, C_FF, C_MI
    type t_Vei is record
    (
        Ef number, -- Entradas detectadas pelo fisco
        Ec number, -- Entradas declaradas pelo grupo
        Sf number, -- Saídas detectadas pelo fisco
        Sc number, -- Saídas declaradas pelo grupo
        Vif number, -- Valor do imposto detectado pelo fisco
        Vic number, -- Valor do imposto declarado pelo contribuinte
        PSD number, -- Probabilidade do grupo de ser avaliado pelo fisco
        Vs number  -- Valor da sanção aplicada pelo fisco no caso de
    descoberta de fraude pelo grupo
    );

    -- Vetor com histórico de estados internos do contribuinte
    type t_Vei_Array is table of t_Vei index by binary_integer;

    -- Vetor do erro dos atributos previstos
    type t_Erro is record
    (
        err_Ef number, -- Erro para atributo Ef
        err_Ec number, -- Erro para atributo Ec
        err_Sf number, -- Erro para atributo Sf
        err_Sc number, -- Erro para atributo Sc
        err_Vif number, -- Erro para atributo Vif
        err_Vic number, -- Erro para atributo Vic
        err_PSD number, -- Erro para atributo PSD
        err_Vs number  -- Erro para atributo Vs
    );

    -- Vetor com histórico dos erro dos atributos previstos
    type t_Erro_Array is table of t_erro index by binary_integer;

    -- Prevê o comportamento para um contribuinte, retornando
    -- um registro com os valores que definem o estado interno deste

```

```

contribuinte
  function fnc_comportamento_previsto_ie
  (
    p_perodo_inicial number, -- Período inicial dos dados históricos
    (formato YYYYMM)
    p_perodo_final number, -- Período final dos dados históricos
    (formato YYYYMM)
    p_perodo_previsao number, -- Período cujo comportamento será
previsto
    p_metodo number, -- Método de forecasting usado
    p_ie number -- Inscrição estadual cujo comportamento
será previsto
  ) return t_Vei;

  -- Prevê o comportamento para um grupo de contribuintes, retornando
  -- um registro com os valores que definem o estado interno deste
grupo.
  -- Usa a fnc_comportamento_previsto_ie para compor seus valores
  function fnc_comportamento_previsto_grupo
  (
    p_perodo_inicial number, -- Período inicial dos dados históricos
    (formato YYYYMM)
    p_perodo_final number, -- Período final dos dados históricos
    (formato YYYYMM)
    p_perodo_previsao number, -- Período cujo comportamento será
previsto
    p_metodo number, -- Método de forecasting usado
    p_nivel_agregacao number -- Tipo do grupo cujo comportamento será
previsto:
    -- C_SG, C_SE, C_GE, C_AE, C_FF ou C_MI
  ) return t_Vei;

  -- Consulta o comportamento real de um contribuinte, retornando
  -- um vetor com os registros que definem os estados interno deste
contribuinte
  -- para cada mês
  function fnc_comportamento_real_ie
  (
    p_perodo_inicial number, -- Período inicial dos dados históricos
    (formato YYYYMM)
    p_perodo_final number, -- Período final dos dados históricos
    (formato YYYYMM)
    p_ie number -- Inscrição estadual cujo comportamento
real será retornado
  ) return t_Vei_Array;

  -- Consulta o comportamento real de um grupo, retornando
  -- um vetor com os registros que definem os estados internos deste
grupo
  -- para cada mês. Usa a fnc_comportamento_real_ie para compor seus
valores
  function fnc_comportamento_real_grupo
  (
    p_perodo_inicial number, -- Período inicial dos dados históricos
    (formato YYYYMM)
    p_perodo_final number, -- Período final dos dados históricos
    (formato YYYYMM)

```

```
    p_nivel_agregacao number, -- Tipo do grupo cujo comportamento será  
retornado:                                -- C_SG, C_SE, C_GE, C_AE, C_FF ou C_MI  
  ) return t_Vei_Array;  
  
  -- Retorna um vetor com os erros percebidos  
  -- entre o comportamento previsto e o comportamento real de um grupo  
  function fnc_erro_previsao_periodo  
  (  
    p_Vei_previsto t_Vei_Array,  
    p_Vei_real t_Vei_Array  
  ) return t_Erro_Array;  
  
end;
```

ANEXO II – RESULTADOS

Legenda:

SG – Supergrupo de contribuintes (todos os contribuintes)

SE – Setor econômico (primário, secundário e terciário)

GE – Grupo econômico (atacadista, varejista, indústria)

AE – Atividade econômica (tabela nacional do CNAE)

FF – Faixa de faturamento (50 faixas de faturamento)

MI – Modelo individual (um modelo por contribuinte)

Série Ef Holt

Mês	SG	SE	GE	AE	FF	MI	SG	SE	GE	AE	FF	MI
	1 passo à frente						6 passos à frente					
Jan	7,09	5,46	5,79	5,54	5,69	5,10	33,74	31,15	30,39	31,01	31,00	27,30
Fev	5,57	6,12	4,34	4,72	4,64	4,88	31,54	35,30	34,41	30,82	33,15	29,38
Mar	7,04	7,38	6,84	7,14	4,88	6,41	31,95	33,76	32,27	30,19	33,96	29,44
Abr	5,41	7,26	5,90	5,55	6,46	6,56	34,11	30,27	30,71	29,05	30,41	28,61
Mai	7,49	6,75	8,71	5,55	6,37	3,72	32,44	30,04	31,87	31,89	30,93	26,50
Jun	5,74	5,65	4,89	3,92	6,49	6,26	30,73	30,19	29,82	25,53	29,93	32,44
Jul	7,11	4,99	7,10	7,52	7,05	7,54	31,68	33,80	32,78	26,68	32,84	29,16
Ago	9,46	4,88	5,73	7,00	8,65	9,59	31,93	29,95	29,25	30,03	30,31	26,76
Set	7,25	5,16	6,89	5,67	5,58	7,00	33,08	33,50	34,17	29,78	32,50	26,88
Out	4,11	5,64	6,55	6,54	6,74	4,54	32,45	34,77	33,54	26,90	34,86	30,79
Nov	6,28	7,64	6,28	5,88	5,87	3,84	33,16	31,44	32,22	27,98	31,17	26,18
Dez	5,21	7,22	5,48	4,45	6,72	5,51	31,39	32,64	31,57	27,32	32,97	27,80
Média	6,48	6,18	6,21	5,79	6,26	5,91	32,35	32,23	31,92	28,93	32,00	28,44

Série Ef Neural

Mês	SG	SE	GE	AE	FF	MI	SG	SE	GE	AE	FF	MI
	1 passo à frente						6 passos à frente					
Jan	7,06	7,43	6,43	6,49	6,58	6,81	25,41	25,57	26,17	25,37	25,85	26,47
Fev	7,34	6,60	5,79	6,50	6,58	9,28	27,92	26,58	25,83	24,73	27,43	27,21
Mar	6,35	6,80	6,51	6,21	6,59	8,74	24,26	24,76	25,24	25,68	25,22	24,44
Abr	6,69	6,00	5,78	5,84	6,57	6,23	25,18	24,82	25,35	24,94	24,98	25,15
Mai	7,05	7,16	6,95	4,33	6,62	3,57	25,21	25,86	24,66	24,69	26,38	25,22
Jun	6,35	5,39	5,39	4,70	5,37	5,33	31,67	31,29	31,57	31,54	31,66	31,04
Jul	7,09	5,71	5,81	4,74	5,82	3,02	25,45	26,72	26,73	27,70	26,84	25,85
Ago	5,29	6,50	5,34	6,19	5,64	3,90	23,19	22,62	22,14	21,53	22,36	22,56
Set	5,58	5,10	4,90	5,52	4,15	3,25	22,61	24,41	24,08	24,36	23,31	23,45
Out	6,47	5,80	5,96	5,20	5,91	4,44	30,55	29,77	30,46	30,66	30,39	30,62
Nov	3,75	3,70	4,48	5,19	4,14	5,45	25,53	24,97	24,54	24,70	25,11	25,45
Dez	5,87	5,92	6,30	5,09	6,04	4,46	25,78	25,63	25,38	25,47	25,98	25,53
Média	6,24	6,01	5,80	5,50	5,83	5,38	26,06	26,08	26,01	25,95	26,29	26,08

Série Ec Holt

Mês	SG	SE	GE	AE	FF	MI	SG	SE	GE	AE	FF	MI
-----	----	----	----	----	----	----	----	----	----	----	----	----

	1 passo à frente						6 passos à frente					
	Jan	Fev	Mar	Abr	Mai	Jun	Jan	Fev	Mar	Abr	Mai	Jun
Jan	7,17	6,27	6,06	4,45	4,96	4,59	32,68	30,98	30,72	30,91	31,21	26,90
Fev	5,87	5,20	4,80	4,90	4,74	5,30	32,19	34,08	34,66	29,78	33,74	30,33
Mar	6,57	7,78	6,74	6,84	4,75	6,89	32,07	33,23	33,48	29,67	33,15	30,22
Abr	5,30	6,87	5,68	6,58	7,19	7,82	34,68	30,17	30,83	28,73	30,84	27,64
Mai	7,05	6,78	8,29	6,09	5,18	2,99	33,23	30,87	31,30	31,62	30,97	26,45
Jun	5,91	6,34	5,32	3,37	6,32	4,92	31,36	29,92	29,33	26,35	29,79	32,42
Jul	7,20	4,58	6,69	6,58	7,60	7,24	31,93	32,61	32,87	26,69	33,33	28,35
Ago	9,15	4,52	5,22	7,08	8,11	9,39	31,40	30,26	30,52	30,37	30,21	27,70
Set	7,27	6,28	7,37	5,95	5,05	8,07	32,40	33,85	34,01	29,37	32,26	27,21
Out	3,79	5,88	6,60	6,95	7,00	4,03	31,87	33,95	33,63	27,26	35,68	30,60
Nov	6,05	7,07	5,45	5,77	5,15	3,90	32,83	31,48	31,90	27,40	30,48	25,83
Dez	5,45	6,99	5,18	4,80	6,99	4,89	30,87	32,89	31,46	27,98	33,23	27,76
Média	6,40	6,21	6,12	5,78	6,09	5,84	32,29	32,02	32,06	28,84	32,07	28,45

Série Ec Neural

Mês	SG	SE	GE	AE	FF	MI	SG	SE	GE	AE	FF	MI
	1 passo à frente						6 passos à frente					
Jan	7,80	7,53	6,29	7,08	7,18	6,79	26,41	25,62	26,34	25,55	25,98	26,54
Fev	6,55	6,51	6,86	6,63	7,23	9,07	27,42	27,20	26,44	24,66	27,31	26,49
Mar	5,20	7,18	5,87	6,75	6,00	8,91	23,81	24,86	25,70	24,92	25,84	24,76
Abr	6,51	6,27	6,11	6,20	6,79	5,82	24,89	25,20	24,58	24,70	24,38	25,04
Mai	7,05	6,51	6,96	4,62	6,69	3,50	25,04	24,73	23,95	25,38	25,76	25,89
Jun	5,82	6,12	5,26	4,94	5,70	5,41	32,24	31,48	31,87	31,88	32,13	31,28
Jul	6,82	6,55	6,26	6,35	6,83	3,59	25,96	27,50	27,72	27,45	26,93	26,62
Ago	5,45	6,25	5,40	5,75	5,35	4,68	22,41	22,16	22,53	22,02	22,36	22,21
Set	5,48	5,70	5,35	4,24	4,17	2,94	22,76	25,71	24,65	23,96	23,25	22,97
Out	5,75	4,90	6,28	5,02	5,43	4,47	30,64	28,82	29,80	31,14	31,14	29,69
Nov	4,02	3,93	4,66	5,03	4,75	5,96	25,77	24,76	24,84	23,77	25,46	25,57
Dez	6,19	6,02	5,10	5,77	6,42	5,14	25,07	25,85	25,53	25,05	25,33	25,76
Média	6,05	6,12	5,87	5,70	6,05	5,52	26,03	26,16	26,16	25,87	26,32	26,07

Série Sf Holt

Mês	SG	SE	GE	AE	FF	MI	SG	SE	GE	AE	FF	MI
	1 passo à frente						6 passos à frente					
Jan	7,26	6,47	6,17	5,24	5,83	5,26	33,46	30,66	30,70	31,51	32,08	27,44
Fev	6,17	6,30	4,79	4,83	4,11	4,71	31,45	33,65	35,13	29,65	34,09	30,06
Mar	7,21	6,88	6,06	6,76	5,33	6,70	32,05	33,64	33,52	29,32	33,24	30,72
Abr	5,49	7,90	6,19	6,61	6,09	7,70	33,79	30,20	30,75	28,70	30,47	27,37
Mai	6,39	7,08	8,38	5,20	6,10	3,79	32,96	31,19	31,70	31,11	31,17	26,36
Jun	6,60	5,58	6,39	3,67	6,44	5,70	30,74	29,36	29,99	25,79	29,15	32,90
Jul	7,14	5,49	6,66	7,34	7,10	7,88	32,33	33,59	32,09	25,96	33,25	28,28
Ago	8,18	5,12	5,25	7,14	8,66	9,68	31,17	30,13	29,54	30,36	29,64	27,95
Set	6,78	4,72	6,70	6,65	6,18	6,96	32,99	33,36	34,81	29,61	33,24	26,79
Out	4,08	5,68	6,67	6,08	6,11	4,59	31,64	34,11	33,65	27,11	34,88	30,55
Nov	5,71	7,12	5,63	4,94	6,35	3,96	33,22	30,81	31,47	28,34	31,31	26,01
Dez	4,60	6,76	4,72	5,59	5,88	4,92	30,81	32,68	31,67	28,07	33,15	28,95
Média	6,30	6,26	6,13	5,84	6,18	5,99	32,22	31,95	32,08	28,79	32,14	28,62

Série Sf Neural

Mês	SG	SE	GE	AE	FF	MI	SG	SE	GE	AE	FF	MI
	1 passo à frente						6 passos à frente					
Jan	7,17	7,56	6,26	6,29	7,02	7,21	26,28	25,75	25,85	26,03	25,55	25,95
Fev	7,22	5,84	6,27	7,25	6,83	9,92	27,15	26,25	25,41	24,61	27,59	27,35

Mar	6,09	6,68	6,04	6,33	6,30	9,72	24,18	24,35	25,24	25,12	25,24	24,56
Abr	7,18	6,58	5,90	6,85	6,46	5,28	25,01	24,71	24,13	24,59	24,78	25,06
Mai	6,59	7,47	6,87	4,03	6,58	3,78	25,67	25,41	24,76	24,91	26,10	25,89
Jun	5,96	6,52	6,24	5,13	6,19	5,37	32,59	30,76	31,96	32,10	32,01	30,84
Jul	6,62	6,60	5,12	4,76	6,13	3,36	25,55	26,71	26,78	26,51	27,24	26,84
Ago	4,62	5,80	6,30	6,14	6,31	4,28	23,07	22,03	22,41	21,23	22,35	22,10
Set	4,79	5,36	4,47	4,70	4,67	3,22	24,14	24,88	23,72	25,14	24,07	23,23
Out	6,86	5,04	6,76	5,40	6,71	4,39	30,60	29,57	29,46	31,19	31,01	29,58
Nov	4,14	4,11	4,56	5,26	4,18	4,31	26,08	25,35	25,08	24,35	25,13	25,15
Dez	6,83	5,64	6,77	4,77	6,36	5,44	25,16	25,74	25,77	24,95	26,32	24,73
Média	6,17	6,10	5,96	5,58	6,15	5,52	26,29	25,96	25,88	25,89	26,45	25,94

Série Sc Holt

Mês	SG	SE	GE	AE	FF	MI	SG	SE	GE	AE	FF	MI
	1 passo à frente						6 passos à frente					
Jan	7,54	6,17	5,73	5,84	5,10	5,38	32,92	30,57	30,46	31,00	30,85	27,57
Fev	5,29	5,53	4,30	5,39	5,50	5,07	31,75	33,89	35,20	29,95	34,17	30,37
Mar	6,14	6,46	5,54	6,88	5,72	6,69	31,64	34,19	33,41	29,71	33,48	30,13
Abr	4,79	7,27	6,88	6,16	6,35	6,30	34,17	30,38	31,27	29,57	31,02	27,46
Mai	7,42	7,08	8,41	5,68	5,57	3,78	33,42	31,06	31,66	31,00	30,90	25,80
Jun	5,97	6,17	5,53	3,39	5,79	5,94	30,75	29,50	29,23	25,85	28,87	32,54
Jul	7,62	5,30	7,03	6,78	7,23	7,90	31,70	32,58	33,25	26,81	34,23	28,30
Ago	9,38	5,25	5,55	7,18	9,03	9,80	31,67	30,66	30,25	30,42	30,41	26,64
Set	6,48	5,53	7,43	5,58	6,11	7,93	32,72	33,43	34,21	29,82	32,95	26,88
Out	4,11	6,09	5,67	7,03	7,03	4,07	32,12	33,55	33,75	27,19	35,00	30,34
Nov	6,12	7,75	6,55	5,50	6,05	4,54	33,44	31,57	32,48	27,73	30,92	27,02
Dez	4,30	6,32	4,86	5,37	6,55	4,69	31,17	32,98	31,17	27,98	32,43	29,01
Média	6,26	6,24	6,12	5,90	6,33	6,01	32,29	32,03	32,19	28,92	32,10	28,50

Série Sc Neural

Mês	SG	SE	GE	AE	FF	MI	SG	SE	GE	AE	FF	MI
	1 passo à frente						6 passos à frente					
Jan	6,83	7,91	6,64	7,05	6,66	6,99	26,08	25,01	25,55	25,24	25,34	25,94
Fev	7,00	6,34	5,49	6,77	7,12	10,16	28,21	26,62	26,09	24,53	27,24	26,94
Mar	5,96	6,14	6,33	6,66	6,87	8,85	24,30	25,07	25,27	26,18	25,08	24,33
Abr	6,43	6,05	6,24	7,19	6,82	5,31	25,24	25,13	25,22	23,82	24,30	24,58
Mai	7,16	7,25	6,86	3,95	6,39	3,42	25,25	25,00	25,47	24,72	26,16	25,18
Jun	5,18	5,74	5,37	5,18	6,18	4,46	32,46	31,80	30,56	31,43	31,46	31,49
Jul	7,00	6,03	6,33	5,89	5,66	2,90	26,58	26,48	27,56	27,55	26,76	26,79
Ago	5,93	5,90	6,37	5,76	5,68	3,72	22,63	22,87	21,83	21,67	22,53	22,10
Set	5,28	4,94	4,55	4,89	4,68	3,87	23,53	25,24	24,28	24,88	23,86	23,07
Out	6,27	5,35	5,44	4,84	5,99	4,72	30,83	29,24	30,21	30,72	29,96	30,07
Nov	4,90	5,05	5,01	4,50	4,95	5,45	25,27	24,94	24,60	23,66	24,69	25,45
Dez	6,63	6,40	5,61	4,46	6,25	5,24	24,65	25,37	25,38	25,49	25,41	25,48
Média	6,21	6,09	5,85	5,60	6,10	5,42	26,25	26,07	26,00	25,82	26,07	25,95

Série Vif Holt

Mês	SG	SE	GE	AE	FF	MI	SG	SE	GE	AE	FF	MI
	1 passo à frente						6 passos à frente					
Jan	6,83	6,26	5,94	5,63	6,10	4,70	33,10	31,32	31,24	30,26	31,42	26,91
Fev	5,91	5,89	4,18	5,55	4,24	4,88	31,99	34,99	34,91	30,83	33,89	30,04
Mar	6,90	6,52	5,69	6,33	4,88	7,52	32,39	34,04	32,81	29,34	34,56	30,27
Abr	4,70	7,78	6,02	6,57	6,79	6,69	34,61	30,31	30,54	28,18	31,55	28,58
Mai	7,79	6,08	8,04	6,22	6,25	4,28	32,61	30,35	31,34	31,57	30,54	26,26

Jun	6,68	5,14	5,68	3,77	6,17	5,85	30,90	29,59	29,53	25,87	30,10	32,37
Jul	8,07	5,59	6,80	6,69	7,09	7,68	31,56	32,75	32,43	26,23	32,87	28,53
Ago	8,64	5,16	6,06	7,15	8,44	9,64	31,56	29,82	29,61	29,81	30,17	26,43
Set	7,23	4,75	8,03	6,35	6,32	7,24	32,12	32,65	33,56	29,48	33,50	26,28
Out	4,60	5,21	5,76	7,01	6,73	4,90	32,00	34,65	33,40	26,51	35,59	30,61
Nov	5,05	7,95	6,26	5,52	6,43	3,75	33,41	30,70	31,60	27,27	31,26	26,12
Dez	4,99	6,64	4,30	5,24	7,04	4,01	31,44	32,58	31,45	28,23	32,97	28,42
Média	6,45	6,08	6,06	6,00	6,37	5,93	32,31	31,98	31,87	28,63	32,37	28,40

Série Vif Neural

Mês	SG	SE	GE	AE	FF	MI	SG	SE	GE	AE	FF	MI
	1 passo à frente						6 passos à frente					
Jan	6,68	7,21	6,49	7,78	7,05	5,94	26,01	25,77	25,63	25,84	25,15	25,33
Fev	7,02	6,02	5,41	7,12	7,40	9,38	28,48	26,47	26,56	24,69	27,35	26,64
Mar	6,09	7,04	7,21	6,90	7,29	9,44	24,01	24,82	24,70	25,25	25,12	24,58
Abr	6,06	6,74	5,52	7,08	6,51	5,70	25,66	25,44	24,81	24,56	24,11	25,21
Mai	7,06	7,56	6,74	3,43	7,35	4,42	25,18	24,56	24,31	24,42	25,88	25,12
Jun	5,65	5,66	5,36	4,80	5,48	4,49	32,23	31,06	31,21	31,16	32,01	30,72
Jul	6,76	5,94	5,90	5,54	5,89	2,83	26,42	27,30	27,97	26,58	26,34	26,81
Ago	4,81	6,27	6,22	6,52	6,78	4,65	22,62	21,97	22,62	22,36	22,29	22,74
Set	5,29	4,52	4,06	4,66	4,51	3,31	23,73	24,44	24,24	24,63	23,44	22,59
Out	6,50	5,45	6,28	4,53	6,50	4,03	29,78	29,89	29,83	30,68	29,41	29,63
Nov	3,90	3,88	4,57	4,52	3,94	5,16	25,38	25,10	24,40	24,04	25,50	25,98
Dez	5,57	6,80	6,00	5,55	5,28	4,26	25,15	26,23	25,95	25,80	25,06	24,61
Média	5,95	6,09	5,81	5,70	6,17	5,30	26,22	26,09	26,02	25,83	25,97	25,83

Série Vic Holt

Mês	SG	SE	GE	AE	FF	MI	SG	SE	GE	AE	FF	MI
	1 passo à frente						6 passos à frente					
Jan	6,89	6,15	5,99	5,19	5,65	4,98	33,12	31,06	30,70	30,82	31,49	27,55
Fev	5,74	5,81	4,71	4,98	4,54	4,64	31,40	34,40	34,85	30,30	33,92	30,16
Mar	6,85	6,92	5,94	7,08	4,93	6,97	32,39	33,71	33,23	29,77	33,92	30,03
Abr	5,35	7,80	6,19	6,37	6,51	7,17	34,30	30,58	30,99	29,00	31,05	28,19
Mai	7,17	6,53	8,09	5,88	5,94	3,78	32,67	30,82	31,14	31,35	30,94	26,32
Jun	5,88	5,57	5,59	3,76	6,38	5,88	31,02	29,75	29,86	26,25	29,66	32,29
Jul	7,80	5,33	6,48	7,04	7,29	7,80	32,16	33,22	32,89	26,75	33,42	28,73
Ago	8,86	5,15	5,65	7,28	8,60	9,30	31,57	30,43	29,69	29,69	30,23	27,01
Set	7,38	5,48	7,57	6,18	5,80	7,73	32,71	33,26	34,12	29,47	32,69	26,98
Out	4,28	5,82	6,53	6,85	6,55	4,14	32,01	34,27	33,65	27,04	34,84	30,41
Nov	5,52	7,21	5,86	5,50	5,99	3,96	33,38	31,19	32,39	27,65	31,12	26,23
Dez	4,63	6,98	4,93	5,31	6,37	4,63	30,93	32,87	31,88	28,13	33,03	28,31
Média	6,36	6,23	6,13	5,95	6,21	5,92	32,31	32,13	32,12	28,85	32,19	28,52

Série Vic Neural

Mês	SG	SE	GE	AE	FF	MI	SG	SE	GE	AE	FF	MI
	1 passo à frente						6 passos à frente					
Jan	7,28	7,60	6,18	6,84	6,95	6,72	25,90	25,82	25,82	25,67	25,60	25,84
Fev	7,20	6,50	6,19	6,52	6,79	9,85	27,91	26,27	25,96	24,86	27,42	26,72
Mar	6,08	6,62	6,66	6,63	6,63	8,97	24,40	25,11	25,35	25,57	25,27	24,56
Abr	6,38	6,56	5,92	6,51	6,65	5,72	25,55	24,8	24,90	24,07	24,61	24,93
Mai	6,88	6,91	6,73	4,19	6,92	3,67	25,35	25,12	24,79	25,07	25,83	25,45
Jun	6,01	5,92	5,59	4,94	5,72	4,82	32,17	31,29	31,46	31,43	31,86	31,16
Jul	6,77	6,26	5,88	5,54	6,22	3,56	26,23	26,67	27,13	27,23	26,76	26,45

Ago	5,26	6,08	6,1	5,99	6,08	3,91	22,75	22,52	22,40	22,06	22,54	22,32
Set	5,24	5,07	4,53	4,72	4,50	3,51	23,43	24,79	24,25	24,32	23,60	22,74
Out	6,28	5,72	6,2	5,06	6,28	4,75	30,32	29,25	29,97	30,64	30,32	29,86
Nov	4,10	4,42	4,43	4,68	4,25	5,03	25,94	25,38	24,82	24,36	25,38	25,51
Dez	6,23	5,99	5,81	5,1	6,19	4,84	25,28	25,62	25,64	25,81	25,90	25,29
Média	6,14	6,13	5,85	5,56	6,10	5,45	26,27	26,05	26,04	25,92	26,26	25,90

Série PSD Holt

Mês	SG	SE	GE	AE	FF	MI	SG	SE	GE	AE	FF	MI
	1 passo à frente						6 passos à frente					
Jan	7,02	6,47	6,21	5,94	5,59	4,57	33,10	31,38	30,83	30,83	31,76	28,07
Fev	5,24	5,74	5,25	4,47	4,74	5,17	30,71	34,92	35,51	30,44	34,20	29,93
Mar	7,31	7,47	6,50	7,22	5,47	7,54	31,74	34,05	33,73	29,70	33,94	29,84
Abr	4,63	7,44	6,12	6,48	5,68	7,44	33,89	31,11	30,83	29,08	31,18	28,10
Mai	6,69	5,70	8,39	6,58	5,74	3,69	32,13	31,39	31,77	31,30	31,13	26,77
Jun	5,87	5,33	5,35	3,61	6,46	6,63	30,78	29,06	30,35	26,30	29,78	32,47
Jul	7,27	6,03	6,78	7,44	7,44	8,47	31,83	33,24	33,56	26,68	33,00	28,02
Ago	9,09	5,18	5,22	7,96	7,99	9,31	32,01	30,58	29,70	29,64	30,72	26,67
Set	7,50	5,69	8,36	5,89	5,37	7,96	32,72	33,31	34,48	29,67	32,46	27,30
Out	4,53	5,67	6,66	7,09	6,24	3,98	31,45	33,41	33,37	27,01	34,20	30,42
Nov	5,30	7,16	5,89	5,23	6,23	4,18	34,03	31,50	33,19	27,86	31,57	26,43
Dez	5,03	7,55	4,71	5,92	5,94	4,36	30,82	32,65	32,05	27,45	33,46	28,34
Média	6,29	6,29	6,29	6,15	6,07	6,11	32,10	32,22	32,45	28,83	32,29	28,53

Série PSD Neural

Mês	SG	SE	GE	AE	FF	MI	SG	SE	GE	AE	FF	MI
	1 passo à frente						6 passos à frente					
Jan	7,02	7,60	6,13	6,86	7,25	6,25	25,64	25,77	26,17	25,20	24,84	25,43
Fev	7,31	6,21	7,09	6,02	6,59	9,55	27,94	25,58	25,53	24,79	27,19	26,58
Mar	6,47	6,75	6,95	6,73	7,20	9,44	24,22	24,43	25,51	26,20	25,37	24,58
Abr	6,66	6,84	5,65	6,44	6,15	5,94	25,54	25,19	25,09	23,56	23,96	25,01
Mai	6,74	6,76	6,18	3,52	7,20	3,29	25,27	24,65	24,98	25,21	25,85	25,57
Jun	5,55	6,27	5,01	4,43	6,19	4,99	32,15	30,74	30,73	31,27	32,10	31,39
Jul	7,60	6,44	6,24	6,29	5,88	3,86	26,43	26,54	27,56	27,30	26,22	25,93
Ago	6,15	5,96	6,65	6,01	6,02	4,59	23,05	22,10	21,96	21,82	21,88	22,49
Set	5,15	5,14	4,40	4,74	4,07	3,63	23,25	24,66	24,03	24,94	23,99	23,13
Out	6,13	6,34	5,97	5,10	6,00	4,75	30,78	29,33	29,84	30,64	30,32	30,14
Nov	4,26	4,08	3,83	4,93	4,30	5,14	25,58	25,71	24,84	24,90	26,07	25,91
Dez	7,11	6,14	4,85	5,33	6,00	4,65	25,60	26,02	25,86	26,02	26,00	24,48
Média	6,35	6,21	5,75	5,53	6,07	5,51	26,29	25,89	26,01	25,99	26,15	25,89

Série Vs Holt

Mês	SG	SE	GE	AE	FF	MI	SG	SE	GE	AE	FF	MI
	1 passo à frente						6 passos à frente					
Jan	7,36	5,30	5,44	4,83	5,53	4,82	33,25	30,82	29,88	31,30	31,19	26,76
Fev	5,49	6,00	4,44	5,32	4,66	5,28	31,73	34,27	35,02	31,10	34,48	29,39
Mar	6,93	7,33	6,47	7,18	4,49	7,37	31,83	33,96	33,40	28,96	34,59	30,47
Abr	5,16	7,94	5,55	6,00	6,90	7,24	34,03	30,65	30,91	29,29	31,13	28,77
Mai	6,45	6,79	7,72	6,16	5,78	3,40	33,11	31,16	31,57	31,34	30,25	25,66
Jun	6,18	5,88	5,47	4,09	6,54	5,48	31,05	29,45	29,48	26,10	29,97	31,95
Jul	7,97	5,05	6,65	6,95	7,27	8,06	32,61	33,69	33,72	26,61	32,71	28,73
Ago	8,59	5,46	6,08	6,96	9,19	8,71	31,67	30,46	30,37	29,18	30,21	26,89
Set	7,15	5,80	7,55	6,57	6,15	7,35	33,01	32,43	34,99	28,90	32,40	26,77
Out	3,75	6,15	5,81	6,92	6,10	3,90	32,45	34,80	33,57	27,29	34,67	30,41

Nov	6,06	7,73	5,64	5,31	6,13	4,08	32,89	31,19	31,66	27,58	31,21	25,30
Dez	5,07	6,94	5,33	4,80	6,68	4,26	31,03	33,46	32,01	28,96	33,45	28,63
Média	6,35	6,36	6,01	5,92	6,28	5,83	32,39	32,19	32,22	28,88	32,19	28,31

Série Vs Neural

Mês	SG	SE	GE	AE	FF	MI	SG	SE	GE	AE	FF	MI
	1 passo à frente						6 passos à frente					
Jan	7,14	8,07	5,72	6,58	7,28	6,78	25,31	26,44	25,82	26,07	24,90	25,35
Fev	7,50	6,81	6,81	6,72	7,04	9,56	27,26	26,09	25,85	24,25	27,50	26,48
Mar	6,50	6,84	5,98	6,30	7,22	8,84	24,42	25,20	25,75	25,71	24,75	23,90
Abr	6,46	6,00	6,16	6,54	7,01	5,99	25,19	25,33	25,58	24,98	24,36	25,29
Mai	6,47	7,18	6,73	4,01	7,40	3,46	24,96	25,43	24,76	25,56	25,23	25,37
Jun	6,58	6,54	5,96	5,47	5,74	4,51	32,54	31,68	31,49	31,06	31,92	30,99
Jul	5,91	6,09	5,72	5,99	6,06	3,42	26,51	27,18	26,88	27,20	27,08	26,28
Ago	6,08	5,27	5,32	5,73	6,01	3,67	23,05	22,76	23,11	22,87	22,54	22,07
Set	5,61	5,79	5,34	4,83	4,32	3,37	23,90	24,33	23,96	23,90	22,74	23,16
Out	5,50	6,20	5,84	4,83	6,32	5,08	29,89	29,68	29,99	31,49	30,59	29,62
Nov	3,67	4,62	3,92	4,22	4,27	4,39	26,51	25,20	24,46	24,20	24,95	25,09
Dez	5,83	6,01	5,24	5,00	6,38	4,68	25,46	24,95	25,56	26,30	25,86	25,57
Média	6,10	6,28	5,73	5,52	6,25	5,31	26,25	26,19	26,10	26,13	26,03	25,76