



UNILASALLE
CENTRO UNIVERSITÁRIO LA SALLE



TIAGO LOCK MARTINS

**CLASSIFICAÇÃO E CATEGORIZAÇÃO DO PROCESSO DE GERENCIAMENTO
DE INCIDENTES BASEADO EM ITIL COM O USO DE TÉCNICAS DE
INTELIGÊNCIA ARTIFICIAL**

CANOAS, 2011

TIAGO LOCK MARTINS

**CLASSIFICAÇÃO E CATEGORIZAÇÃO DO PROCESSO DE GERENCIAMENTO
DE INCIDENTES BASEADO EM ITIL COM O USO DE TÉCNICAS DE
INTELIGÊNCIA ARTIFICIAL**

Trabalho de conclusão apresentado para a banca examinadora do curso de Ciência de Computação do Centro Universitário La Salle – Unilasalle, como exigência parcial para a obtenção do grau de Bacharel em Ciência de Computação.

Orientadora: Prof^a. Dra. Patrícia Kayser Vargas Mangan

CANOAS, 2011

TIAGO LOCK MARTINS

**CLASSIFICAÇÃO E CATEGORIZAÇÃO DO PROCESSO DE GERENCIAMENTO
DE INCIDENTES BASEADO EM ITIL COM O USO DE TÉCNICAS DE
INTELIGÊNCIA ARTIFICIAL**

Trabalho de conclusão apresentado para a banca examinadora do curso de Ciência de Computação do Centro Universitário La Salle – Unilasalle, como exigência parcial para a obtenção do grau de Bacharel em Ciência de Computação.

Aprovado pela banca examinadora em 13 de dezembro de 2011.

BANCA EXAMINADORA:

Prof.^a Dra. Patrícia Kayser Vargas Mangan

Prof. Me. Roberto Petry

Prof. Me. Valderi Reis Quietinho Leithardt

CANOAS, 2011

À minha linda esposa que esteve sempre ao meu lado incentivando e compartilhando meu crescimento em todos os sentidos.

À minha mãe que me direcionou para o caminho correto do conhecimento, dignidade e carinho pelas pessoas.

AGRADECIMENTOS

Agradeço a professora Dr. Patrícia Kayser Vargas Mangan, pela orientação, amizade, paciência e apoio que foram imprescindíveis para o desenvolvimento não somente deste trabalho, mas também para uma grande parte de meu crescimento acadêmico. Ao Sr. Lauro da Silva Martins, pois sem ele não teria plantado a semente que germinou este trabalho, além da constante parceria e apoio.

RESUMO

Muitas empresas, após implementarem seus processos de gestão de Tecnologia da Informação baseados em ITIL, possuem um certo nível de dificuldade na classificação e categorização dos seus registros podendo causar desvios no fluxo e penalizando o processo com retrabalho e, na maioria das vezes, um maior custo de tempo de solução que impacta diretamente nos indicadores de nível de serviço. Este trabalho pretende responder a seguinte questão: É possível utilizar técnicas e algoritmos de Inteligência Artificial para gerar um indicador de probabilidade de desvios, ou acertos, na classificação e categorização de incidentes, tendo como informação de entrada única e exclusivamente a própria base de classificação e categorização? Objetivando avaliar através de experimentações, bases de dados em uso por duas empresas atualmente, empregando técnicas de Inteligência Artificial específicas: Árvores de Decisão J48 e Redes Neurais *Perceptron* Multicamada. Depois de selecionado o Weka como ferramenta para experimentação, recebidas as bases de dados das empresas, normalizados os dados em banco de dados SQL, codificado os valores e relacionamentos em números primos, definidos os parâmetros de execução de cada algoritmo e suas variações, uma sequência de experimentos foram executados com ambas as bases. Os resultados dos experimentos são muito animadores, além da economia de recursos computacionais e tempo com a codificação dos dados em números primos, foi demonstrado que os dois algoritmos são passíveis de auxiliar a avaliação da classificação e categorização das duas bases de dados utilizadas, principalmente a árvore de decisão, porém a rede neural não é aderente a uma das bases.

Palavras-chave: ITIL, Classificação, Inteligência Artificial, Redes Neurais, Árvores de Decisão, Weka.

ABSTRACT

Many companies, after implement their processes of information technology management based on ITIL, have a certain difficulty level in the classification and categorization of your records and this may cause deviations in stream and penalizing process with rework and, in most cases, a higher cost of solution time that impacts directly on the service level indicators. This work seeks to answer the following question: Is it possible to use techniques and algorithms of Artificial Intelligence to generate an indicator of the likelihood of misappropriation, or hits, in the classification and categorization of incidents, taking as input information solely and exclusively the very basis of the classification and categorization? In order to assess through trials, databases in use by the two companies currently employing Artificial Intelligence techniques: specific Decision trees and *Multi-layer perceptron* neural networks J48. Once selected Weka as tool for experimentation, received the corporate databases, normalized data in SQL database, coded the values and relationships in prime numbers, defined the parameters of implementation of each algorithm and its variations, a sequence of experiments were performed with both bases. The results of the experiments are very encouraging, in addition to the economics of computing resources and time with the encoding of data into primes, it was demonstrated that the two algorithms are likely to assist the evaluation of classification and categorization of the two databases used, mainly the decision tree, but the neural network is not adhering to one of the bases.

Keywords: ITIL, Classification, Artificial Intelligence, Neural Networks, Decision Trees, Weka.

LISTA DE EQUAÇÕES

EQUAÇÃO 1 PROBABILIDADE LOGARÍTMICA BÁSICA	39
EQUAÇÃO 2 FUNÇÃO DE ATIVAÇÃO	42
EQUAÇÃO 3 CÁLCULO DE RETRO PROPAGAÇÃO	45

LISTA DE GRÁFICOS

GRÁFICO 1 TREINAMENTO DA BASE 1 – NÍVEIS OCULTOS: COERÊNCIA X TEMPO	58
GRÁFICO 2 TREINAMENTO DA BASE 1 – ÉPOCAS: COERÊNCIA X TEMPO	58
GRÁFICO 3 TREINAMENTO DA BASE 2 – NÍVEIS OCULTOS: COERÊNCIA X TEMPO	59
GRÁFICO 4 TREINAMENTO DA BASE 2 – ÉPOCAS: COERÊNCIA X TEMPO	59
GRÁFICO 5 TREINAMENTO DA BASE 1 – VIAS: COERÊNCIA X TEMPO	60
GRÁFICO 6 TREINAMENTO DA BASE 2 – VIAS: COERÊNCIA X TEMPO	60
GRÁFICO 7 COMPARATIVO DE RESULTADOS DE COERÊNCIA DOS EXPERIMENTOS.....	62

LISTA DE FIGURAS

FIGURA 1 PROCESSO	17
FIGURA 2 CICLO DE VIDA DE PROCESSO.....	23
FIGURA 3 FLUXO DE TRATAMENTO DE INCIDENTES	28
FIGURA 4 FLUXO DE TRATAMENTO DE REQUISIÇÕES.....	30
FIGURA 5 MODELO PDCA APLICADO À CSI	35
FIGURA 6 MODELO MATEMÁTICO SIMPLES PARA UM NEURÔNIO	42
FIGURA 7 REDE DE <i>PERCEPTRONS</i>	43
FIGURA 8 ORGANIZAÇÃO DAS CAMADAS EM UMA MLP	44
FIGURA 9 BASE 1 – TABELA DE DADOS	53
FIGURA 10 BASE 2 – RELACIONAMENTOS	54

LISTA DE QUADROS

QUADRO 1 EQUILÍBRIO ENTRE SERVIÇOS DE TI X COMPONENTES TECNOLÓGICOS.....	25
QUADRO 2 EQUILÍBRIO ESTABILIDADE X RESPONSABILIDADE	25
QUADRO 3 EQUILÍBRIO QUALIDADE DO SERVIÇO X CUSTO DO SERVIÇO	25
QUADRO 4 EQUILÍBRIO REATIVO X PROATIVO.....	26
QUADRO 5 CARACTERÍSTICAS BÁSICAS DAS BASES DE DADOS DE ESTUDO	51

LISTA DE TABELAS

TABELA 1 RESULTADOS DE TREINAMENTO ALGORITMO DE ÁRVORE DE DECISÃO J48	57
TABELA 2 COMPARATIVO DE RESULTADOS DE COERÊNCIA DOS EXPERIMENTOS.....	61

LISTA DE ABREVIATURAS OU SIGLAS

API	<i>Application Programming Interface</i>
ARFF	<i>Attribute-Relation File Format</i>
CAB	<i>Change Advisory Board</i>
CCM	Conselho de Controle de Mudanças
CMDB	<i>Configuration Management Database</i>
CSI	<i>Continual Service Improvement</i>
CSV	<i>Comma Separated Values</i>
DBEC	Base de Dados de Erros Conhecidos
GB	<i>GigaBytes</i>
GHz	<i>GigaHertz</i>
GNS	Gerenciamento de Nível de Serviço
GSTI	Gerenciamento de Serviços de TI
IA	Inteligência Artificial
IC	Item de Configuração
IT	<i>Information Technology</i>
ITIL	<i>Information Technology Infrastructure Library</i>
ITILv2	Segunda versão da ITIL
ITILv3	Terceira Versão da ITIL
ITSM	<i>Information Technology Service Management</i>
JDBC	<i>Java Database Conector</i>
K+	Potássio
KDD	<i>Knowledge Discovery in Databases</i>
KEDB	<i>Known Error Database</i>
KPI	<i>Key Performance Indicator</i>
MB	<i>MegaBytes</i>
MDC	Máximo Divisor Comum
MLP	<i>Multi-Layer Perceptron</i>
MVS	Máquina de vetor de suporte
Na+	Sódio
PDCA	<i>Plan, Do, Check, Act</i>
RDM	Requisição de Mudança

RFC	<i>Request for Change</i>
SLA	<i>Service Level Agreement</i>
SQL	<i>Structured Query Language</i>
T-SQL	<i>Transact SQL</i>
TCO	<i>Total Cost of Ownership</i>
TI	<i>Tecnologia da Informação</i>
WEKA	<i>Waikato Environment for Knowledge Analysis</i>
XOR	<i>Exclusive or</i>

SUMÁRIO

1	INTRODUÇÃO	14
2	FUNDAMENTAÇÃO TEÓRICA	16
2.1	ITILv2	18
2.1.1	<i>Gerenciamento da Configuração</i>	18
2.1.1.1	Banco de dados de gerenciamento de configuração (CMDB)	18
2.1.2	<i>Gerenciamento de Incidentes</i>	19
2.1.2.1	Incidente	19
2.1.3	<i>Central de Serviços</i>	20
2.1.4	<i>Gerenciamento de Problemas</i>	21
2.1.5	<i>Gerenciamento de Mudanças</i>	21
2.1.6	<i>Gerenciamento de Liberações</i>	22
2.2	ITILv3	22
2.2.1	<i>Ciclo de Vida do Serviço</i>	23
2.2.2	<i>Operação de Serviço</i>	24
2.2.3	<i>Gerenciamento de Incidentes</i>	26
2.2.4	<i>Cumprimento de Requisição</i>	29
2.2.5	<i>Gerenciamento de Eventos</i>	30
2.2.5.1	Evento	31
2.2.6	<i>Gerenciamento de Problemas</i>	31
2.2.7	<i>Gerenciamento de Acesso</i>	31
2.2.8	<i>Funções e papéis na Operação do Serviço</i>	32
2.2.8.1	Central de Serviços (<i>Service Desk</i>)	32
2.2.8.2	Gerenciamento Técnico (<i>Technical Management</i>)	33
2.2.8.3	Gerenciamento de Aplicações (<i>Application management</i>)	33
2.2.8.4	Gerenciamento da Operação de TI (<i>IT Operations Management</i>)	34
2.2.8.5	Melhoria de Serviço Continuada (<i>Continual Service Improvement</i>)	34

2.3	Considerações finais	35
3	APLICAÇÃO DE INTELIGÊNCIA ARTIFICIAL EM DATA MINING.....	36
3.1	Modelos de Aprendizado.....	37
3.1.1	<i>Máxima probabilidade: modelos discretos.....</i>	<i>38</i>
3.2	Árvore de Decisão.....	39
3.2.1	<i>Algoritmos clássicos</i>	<i>40</i>
3.3	Redes Neurais	41
3.3.1	<i>Breve Histórico da Rede Neural Artificial.....</i>	<i>41</i>
3.3.2	<i>Características das Redes Neurais</i>	<i>41</i>
3.3.2.1	<i>Perceptron.....</i>	<i>42</i>
3.3.2.2	<i>Multi-layer perceptron (MLP).....</i>	<i>43</i>
3.4	O Weka.....	46
3.5	Considerações finais	47
4	DESENVOLVIMENTO	48
4.1	Características das bases de dados.....	48
4.2	Escolha dos algoritmos.....	51
4.3	Configurando base de dados.....	51
4.4	Experimentações.....	54
4.5	Análise dos resultados	56
4.6	Considerações finais	62
5	CONCLUSÃO	63
	APÊNDICE A – TEOREMA FUNDAMENTAL DA ARITMÉTICA.....	67
	APÊNDICE B – BASE 1	68
	APÊNDICE C – BASE 2	82

1 INTRODUÇÃO

Muitas empresas, após implementarem seus processos de gestão de Tecnologia da Informação baseados em *Information Technology Infrastructure Library* (ITIL), mesmo utilizando ferramentas de mercado para Gerência de Incidentes no *Service Desk*, possuem um certo nível de dificuldade na classificação e categorização dos seus registros. Problemas na classificação e categorização podem causar desvios no fluxo e penalizando o processo com retrabalho e, na maioria das vezes, um maior custo de tempo de solução que impacta diretamente nos indicadores de nível de serviço.

Neste contexto, formula-se como problema de pesquisa investigado neste trabalho a seguinte questão: É possível utilizar técnicas e algoritmos de Inteligência Artificial para gerar um indicador de probabilidade de desvios, ou acertos, na classificação e categorização de incidentes, tendo como informação de entrada única e exclusivamente a própria base de classificação e categorização?

Este trabalho visa auxiliar a encontrar uma solução otimizada para esta dificuldade através da identificação e experimentação de algoritmos de Inteligência Artificial baseados em árvores de decisão multivaloradas utilizando o algoritmo J48 comparando com os resultados em redes neurais *Multi-layer Perceptron* para a classificação e categorização direcionada.

O foco deste trabalho será as ferramentas para o Gerenciamento de Incidentes e utilizadas pelo *Service Desk*. Atualmente as ferramentas utilizadas mais conhecidas do mercado permitem uma modelagem muito profunda dos processos, registros e fluxos, porém, todas se baseiam em classificação e categorização deste registro. Esta classificação e categorização hoje acontecem em função da percepção do atendente do *Service Desk*, podendo ter grandes variações e exigindo que sejam investidos muitos recursos (tempo e dinheiro) para nivelar todos a cada alteração ou troca de pessoal. Esta situação gera grandes transtornos já que é baseado na classificação que incidem os indicadores de tempo de solução e é na categorização que são definidos os grupos solucionadores de um determinado registro (escalação), portanto podemos ter um registro direcionado erroneamente ou até mesmo o não cumprimento do Acordo de Nível de Serviço (SLA) pela demora na solução, sem falar na dificuldade de controle e avaliação/diagnóstico do processo, já

que pode passar por todo o fluxo e mesmo assim constarem registros idênticos com categorização e/ou classificação diferentes. Existem trabalhos tratando do Gerenciamento de Incidentes utilizando-se de técnicas de inteligência artificial, mas focam em validar a correta categorização de um incidente com base nas informações de entrada, ou apontam para uma possível solução para o incidente, com base na classificação e categorização, a partir da Base de Conhecimento existente.

O principal objetivo deste trabalho é avaliar através de experimentações, bases de dados utilizadas no mercado utilizando técnicas de Inteligência Artificial específicas: Árvores de Decisão J48 e Rede Neurais *Perceptron* Multicamada. Além do objetivo principal, confia-se que este trabalho fornecerá indicações de melhorias para as bases de dados avaliadas, demonstrar como o Weka pode auxiliar estes tipos de experimentações e avaliações, apresentando oportunidades para novos trabalhos, relacionadas ou não com as técnicas ponderadas, a aproveitarem este ambiente.

No próximo capítulo será descrito o contexto ITIL, seus processos e funções. No Capítulo 3 será abordada a área de Inteligência Artificial, incluindo modelos, algoritmos de inferência, Árvores de Decisão, Redes Neurais, *Perceptron* Multicamada, bem como o Weka que é utilizado como API para a modelagem. No Capítulo 4, serão descritas as características das bases de dados, a escolha dos algoritmos, como se dá a configuração das bases de dados para os experimentos, a execução destes experimentos, seus resultados e uma análise crítica dos resultados. O Capítulo 5 finaliza este texto apresentando as considerações finais e trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Com o grande movimento das empresas de médio e grande porte para adaptar seu funcionamento a modelos de Gestão por Processos, as áreas de Tecnologia da Informação (TI) também seguem o mesmo caminho já que fazem parte da organização. No caso de empresas com foco em TI (prestadores de serviço no ramo, por exemplo), isto se torna ainda mais forte em função do controle e da transparência dos serviços.

A ITIL – Information Technology Infrastructure Library – “é uma consolidação de melhores práticas em uma biblioteca que apresenta uma visão consistente e holística do Gerenciamento de Serviços da Tecnologia da Informação, baseada em processos”. (ILLUMNA FOUNDATION CERTIFICATE, 2006)

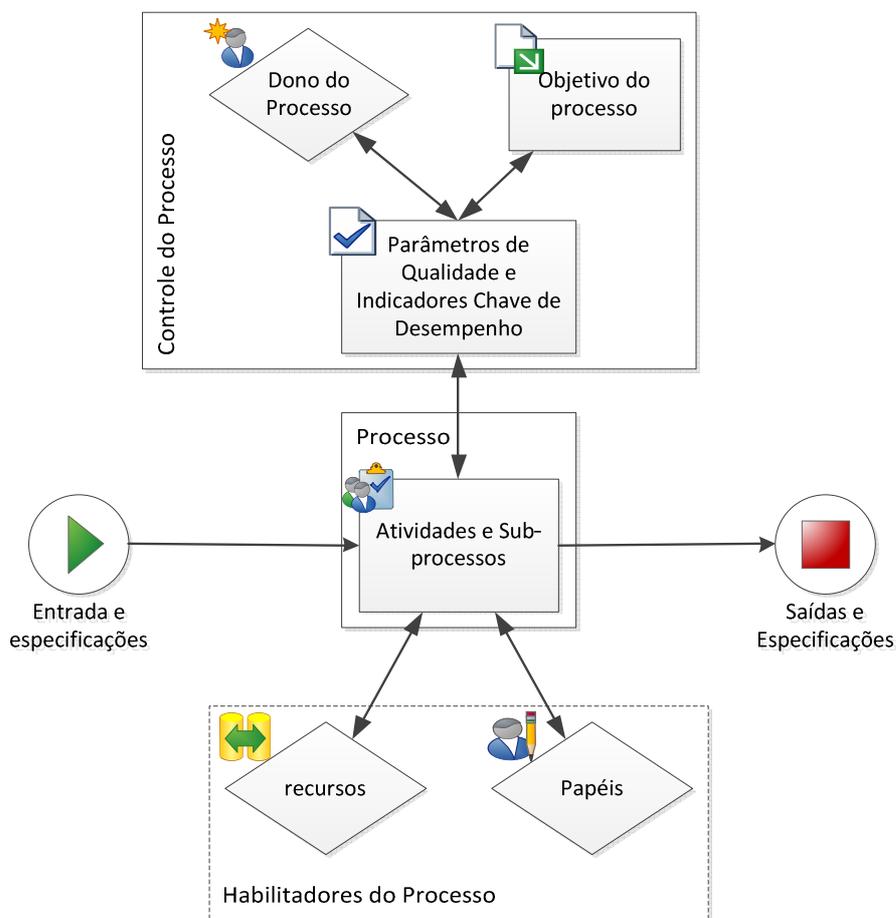
A ITIL é um conjunto de melhores práticas, não uma metodologia, ele sugere onde é possível chegar, para quê, por que, etc, com o objetivo de melhorar o gerenciamento de serviços de TI. Na ITIL um processo é definido por *“uma série concatenada de ações, atividades, mudanças, etc. realizadas por agentes com a intenção de satisfazer uma necessidade ou atingir um objetivo”*. (ILLUMNA FOUNDATION CERTIFICATE, 2006)

Um serviço de TI *“é um conjunto de recursos, TI e não TI, mantidos por um provedor de TI que satisfazem uma ou mais necessidades, suportam os objetivos de negócio do cliente e são percebidos pelo mesmo como um todo coerente”*. (ILLUMNA FOUNDATION CERTIFICATE, 2006)

Na Figura 1 há um modelo de processo utilizado na ITIL. Todo processo possui informações de entrada e/ou disparadores que iniciam o fluxo de trabalho, na expectativa de que o processo gere saídas, isto é, resultados que tragam mais valor para a organização, como relatórios, decisões, procedimentos, revisões do próprio processo, etc. Para gerar estas saídas um processo é composto por várias atividades e subprocessos, como fluxos de trabalho, procedimentos, instruções de trabalho, são executados por unidades organizacionais especializadas em determinado trabalho (Funções) que correspondem a um conjunto de responsabilidades definidas (Papéis) e que, utilizam, combinam e alimentam recursos e habilidades, estes são chamados de Habilitadores do Processo. Para controlar o processo é definido um Dono que representa o processo dentro da

organização, independentemente de onde a tecnologia, serviço e capacidades funcionais estejam alocados e é o principal responsável por garantir que os resultados específicos (objetivos definidos) do processo sejam alcançados, calcado na documentação do processo, suas métricas, parâmetros de qualidade, *feedbacks* e Indicadores Chave de Desempenho (*Key Performance Indicators - KPIs*) para a tomada de decisão.

Figura 1 Processo



Fonte: (ILLUMNA FOUNDATION CERTIFICATE, 2006)

Um serviço de TI “é um conjunto de recursos, TI e não ti, mantidos por um provedor de TI que satisfazem uma ou mais necessidades, suportam os objetivos de negócio do cliente e são percebidos pelo mesmo como um todo coerente”. (ILLUMNA FOUNDATION CERTIFICATE, 2006)

O restante deste capítulo detalha as versões 2 e 3 da ITIL.

2.1 ITILv2

Composto por sete livros em sua segunda versão a ITIL possui dois como núcleo central, o *Service Management* (Gerenciamento de Serviço) composto por *Service Support* (Suporte ao Serviço) e *Service Delivery* (Entrega de Serviço). Estas áreas tratam respectivamente da parte operacional e tática da infraestrutura de TI para suporte ao negócio da empresa, e no caso de empresas de *outsourcing* de TI, a base do próprio negócio. O *Service Support* descreve cinco processos e uma função, Gerenciamento de Configuração, Incidentes, Problemas, Mudanças, Liberações e Central de Serviços (*Service Desk*, função). O *Service Delivery* descreve 5 processos táticos, Gerenciamento de Nível de Serviço, Disponibilidade, Capacidade, Financeiro e Continuidade. O foco principal da segunda versão da ITIL (ITILv2) são os processos de gerenciamento. (ILLUMNA FOUNDATION CERTIFICATE, 2006)

Os conceitos básicos da ITILv2 relacionados diretamente ao trabalho estão apresentados e descritos a seguir.

2.1.1 Gerenciamento da Configuração

O Gerenciamento de configuração é base de todos os outros processos e função, pois fornece as informações sobre o ambiente necessárias para a tomada de decisão e ordenação dos fluxos de trabalho. Ela controla todos os componentes dos serviços de TI, provendo informações atualizadas sobre estes componentes, garantindo a qualidade das informações no Banco de dados de gerenciamento da configuração (CMDB – *Configuration Management Database*). (ILLUMNA FOUNDATION CERTIFICATE, 2006)

2.1.1.1 Banco de dados de gerenciamento de configuração (CMDB)

Conceitual e idealmente seria uma base de dados única contendo todas as informações e dados sobre os componentes e serviços de TI que suportam o negócio, sendo estas informações, apenas o suficiente na visão da organização, para a gestão eficiente e eficaz da infraestrutura de TI. Esta base contém, além dos

itens de configuração (ICs), seus dados, atributos e relacionamentos com outros ICs, como ser utilizado por outro IC, filho de outro IC, parte de outro IC, utiliza outro IC, etc. (ILLUMNA FOUNDATION CERTIFICATE, 2006)

Um item de configuração (IC) é a representação lógica de uma unidade de informação gerenciada. Um IC pode ser uma documentação, um *hardware*, *software*, serviço, grupo de ICs, etc.

2.1.2 Gerenciamento de Incidentes

É o sensor do gerenciamento de serviços de TI, é ele que “nota” algo de anormal no funcionamento do ambiente (Incidente). Sua principal responsabilidade é retornar à operação normal do ambiente. Baseia-se no Gerenciamento de configuração para conhecer as linhas de base, isto é, as configurações e comportamento normal e/ou esperado do ambiente, os ICs e suas relações especificados no Acordo de Nível de Serviço. Encaminham erros identificados para o Gerenciamento de Problemas ou para o Gerenciamento de Mudanças e aplica soluções de contorno para erros conhecidos investigados no Gerenciamento de Problemas. É nele que se baseiam as métricas para o Gerenciamento de Nível de Serviços. (ILLUMNA FOUNDATION CERTIFICATE, 2006)

2.1.2.1 Incidente

Um incidente “*é qualquer evento que afeta ou pode afetar o comportamento normal do ambiente*”, impactando no nível de serviço de TI para o negócio. (ILLUMNA FOUNDATION CERTIFICATE, 2006)

Uma requisição de serviço pode ser tratada como se fosse um incidente caso ele seja simples, em casos mais complexos ela gera uma Mudança e é encaminhado para o Gerenciamento de Mudanças.

Todo incidente deve ser classificado em função de sua prioridade, determinada pelo seu impacto ao negócio e urgência requerida. Isto impacta diretamente no tempo de atendimento que o incidente tem para ser resolvido, pois consta no acordo de nível de serviço o tempo máximo para esta resolução para tal prioridade.

Além da classificação, todo incidente deve possuir uma categorização que irá definir os níveis de suporte e equipe solucionadora, modelando sua escalção. A escalção de um incidente ocorre em dois planos: o plano horizontal é o encaminhamento do incidente para níveis de conhecimento tecnológico maiores, equipes solucionadores com maior especialização e/ou conhecimento do ambiente; o plano vertical é acionado em função da categorização versus a prioridade do incidente, pois este envolve os níveis gerenciais para tomada de decisão quanto a alocação de recursos, ativação de contingências a fim de minimizar o impacto do incidente para o negócio, não permitindo o não cumprimento do SLA ou se este seja inevitável, acionar o plano de comunicação e equipe de gerenciamento de crise. (ILLUMNA FOUNDATION CERTIFICATE, 2006)

2.1.3 Central de Serviços

A Central de Serviços (*Service Desk*) não é um processo, mas sim uma função, isto é, um ator junto aos tantos processos da ITIL, mais diretamente ligado aos Gerenciamentos de Incidentes, Problemas e Mudanças. O principal objetivo desta função é estabelecer um ponto único de contato para os usuários, o registro e início imediato do atendimento aos incidentes, visando uma solução imediata ou sua escalção para outros níveis de suporte. (ILLUMNA FOUNDATION CERTIFICATE, 2006)

Como ponto único de contato, a Central de Serviços é “dona” de todos os incidentes e requisições de serviços registradas, é ela que faz toda a comunicação de andamento e monitoria, bem como pesquisas e *feedbacks* junto aos usuários.

Um dos maiores desafios de uma Central de Serviços é a alta rotatividade de funcionários, o que causa uma perda de ativos intelectuais na função, impactando de forma negativa na eficiência e qualidade da mesma. Outro fator impactante é o treinamento das equipes tanto com relação ao ambiente do cliente, quanto aos procedimentos e conhecimentos técnicos e de negócio. Estes conhecimentos possuem uma complexidade grande e é necessário certo tempo para atingir a eficiência e qualidades esperadas, o que pode acarretar um descrédito da função por não atender as expectativas dos usuários.

2.1.4 Gerenciamento de Problemas

O Gerenciamento de Problemas visa minimizar os impactos gerados por erros dentro da infraestrutura de TI, além de prevenir a recorrência de incidentes causados por estes erros. Seu principal objetivo é identificar a “causa-raiz” dos erros, isto é, a fonte do comportamento não desejado, a qual, sendo solucionada, não gerará mais incidente.

O escopo do Gerenciamento de Problemas vai da identificação de um problema, investigação e diagnóstico de sua causa raiz, elaboração de solução de contorno e/ou definitiva. Quando uma solução de contorno é gerada, o problema passa ser chamado de erro conhecido, pois agora o *Service Desk* tem como resolver os incidentes originados por este erro. O Gerenciamento de Problemas encaminha soluções definitivas para o Gerenciamento de Mudanças planejarem as execuções destas soluções através de Requisições de Mudanças (RDMs), impedindo que novos incidentes sejam gerados por este erro. (ILLUMNA FOUNDATION CERTIFICATE, 2006)

2.1.5 Gerenciamento de Mudanças

O processo de Gerenciamento de Mudanças visa aplicar métodos e técnicas padronizadas para tratar todas as alterações na infraestrutura de TI, assim evitando incidentes e problemas causados no ambiente por estas mudanças. Além destes, o processo faz uma avaliação de impacto, risco e custo das mudanças com visão de Negócio, privilegiando as mudanças que estejam de acordo com a estratégia de Negócio e apresentem um maior retorno para o mesmo. (ILLUMNA FOUNDATION CERTIFICATE, 2006)

Segundo Illumna Foundation Certificate,(2006) *“Uma mudança é toda e qualquer alteração que resulta em uma nova situação de um ou mais Itens de Configuração da infraestrutura de TI”*.

O Conselho de Controle de Mudanças (CCM ou, em inglês, *Change Advisory Board – CAB*) é um grupo de pessoas responsáveis por aprovar mudanças, sua priorização e avaliar a aderência da execução da mesma ao processo. Esta equipe deve ter a capacitação técnica e visão de Negócio necessária para tal, bem como a

autonomia e autoridade para a tomada de decisões relativas às mudanças. (ILLUMNA FOUNDATION CERTIFICATE, 2006)

2.1.6 Gerenciamento de Liberações

O Gerenciamento de Liberações se preocupa em executar mudanças em serviços de TI, sendo responsável por planejar, desenhar, construir, comunicar, configurar e testar conjuntos de componentes para o ambiente operacional de TI. É no Gerenciamento de Liberações em que os Problemas são efetivamente resolvidos, e é este processo que serve de entrada para o Gerenciamento de Configuração, pois são as liberações que alteram ICs. (ILLUMNA FOUNDATION CERTIFICATE, 2006)

2.2 ITILv3

A terceira versão da ITIL (ITILv3) trouxe mais sólido o conceito de Ciclo de Vida de Serviço, sendo este composto por cinco estágios: Estratégia, Desenho, Transição, Operação e Melhoria Contínua. Outra grande alteração foi o foco principal que se intensificou no valor para o Cliente e seu Negócio, direcionando-se claramente para a área de *Outsourcing e incorporando* e integrando-se com os avanços de outros modelos como CMMI, Six Sigma, PMBOK, PRINCE2, COBIT, etc.

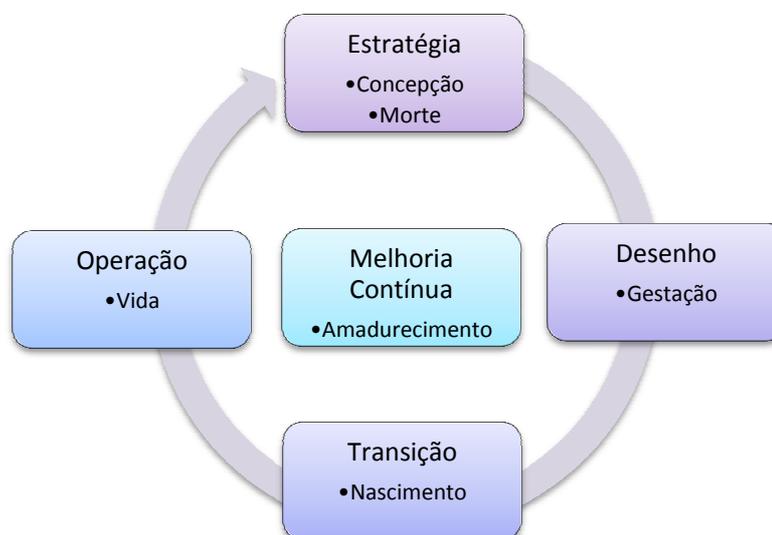
Reduzindo seu volume de sete para cinco livros tendo quatro como núcleo, o *Service Strategy* (Estratégia de Serviço), que foca em criar valor para o Cliente e seu Negócio, ajudando a definir o que deve ser feito, o que pode ser feito, suas prioridades e métricas para garantir o Nível de Serviço. *Service Design* (Desenho do Serviço) responsável por arquitetar o serviço considerando os aspectos de requisitos e soluções, portfólio, tecnologia e arquitetura, processos e métricas de cada serviço novo ou em modificação, focando em reduzir o Custo Total de Propriedade (*Total Cost of Ownership – TCO*). *Service Transition* (Transição de Serviço) visa efetuar a transição de um serviço novo ou modificado à fase de operação, ajustando as expectativas, garantindo que os requisitos sejam obedecidos, controlando riscos, falhas e interrupções, habilitando projetos de mudança, garantindo que o serviço novo ou alterado seja utilizado refletindo as

necessidades de negócio e operação do cliente. O *Service Operation* (Operação de Serviço) seria o equivalente na ITILv3 a uma parte mais operacional do *Service Support* da ITILv2, aprofundando o Gerenciamento de Incidentes e outros e adicionando novos processos e funções além do *Service Desk*. O *Service Operation* também passa a ser onde o balanceamento de conflitos está mais presente, é onde aparecem percepções dos usuários e clientes, equipe de TI, riscos, etc. É nele também onde as operações diárias, a gestão da tecnologia utilizada, monitoramento do ambiente e coordenação de equipes de suporte se apresentam. (ILLUMNA FOUNDATION CERTIFICATE, 2007)

2.2.1 Ciclo de Vida do Serviço

Um dos principais benefícios que a ITILv3 trouxe é a visão de que “*um serviço só atingirá seu objetivo se gerar valor para o Cliente*”, e para tal o processo deve ser acompanhado durante todo seu ciclo de vida segundo Figura 2. (ILLUMNA FOUNDATION CERTIFICATE, 2007)

Figura 2 Ciclo de vida de processo



Fonte: Autoria própria, 2011.

Os objetivos, requisitos, limitações, bem como o valor que o serviço pretende agregar ao cliente e seu negócio são tratados na etapa de Estratégia. É nela que ocorre o “nascimento” de um serviço, bem como a avaliação de sua

descontinuidade, caracterizando sua “morte”. A arquitetura do serviço e seus padrões operacionais são moldados na etapa de Desenho. A etapa de Transição é responsável por colocar em operação o serviço, novo ou alterado, além de efetuar testes, comunicação, alterar Base de Conhecimento, etc.

O serviço gera, de fato, valor para o negócio somente na etapa de Operação, portanto devem-se ter processos muito bem estruturados para garantir a qualidade do serviço e sua aderência às necessidades do cliente, para tanto, evoluindo de forma contínua. Sempre presente durante todo o processo, garantido que todo o aprendizado obtido ao longo de todas as etapas, está a Melhoria Contínua de forma integrada.

2.2.2 *Operação de Serviço*

Coordenando e executando as atividades e processos necessários para entregar e gerenciar os serviços nos níveis acordados para os usuários e clientes, conduzindo, controlando e gerenciando as operações do dia-a-dia, a Operação do Serviço, sem deixar de gerenciar a tecnologia utilizada e envolvida para a entrega e suporte dos serviços, para tanto, deve monitorar o desempenho e identificar métricas e informações para permitir a melhoria contínua dos serviços. É na Operação do Serviço que o cliente percebe o valor. (RHINO CONSULTING, 2010)

A principal característica da Operação do Serviço é o gerenciamento de conflitos, buscando equilíbrio em quatro áreas: Serviços de TI x Componentes Tecnológicos, Estabilidade x Responsividade, Qualidade do Serviço x Custo do Serviço e Reativo x Proativo. Os quadros a seguir apresentam os pontos chave destes conflitos. (ILLUMNA FOUNDATION CERTIFICATE, 2007)

Quadro 1 Equilíbrio entre Serviços de TI x Componentes Tecnológicos

Conflito	Características	Chaves da resolução
<ul style="list-style-type: none"> ✓ Conflito Fundamental ✓ Visão interna (como entregar) ✓ Visão externa (percepção do usuário) ✓ Risco do extremo foco interno <ul style="list-style-type: none"> • Não atingir objetivos de negócio ✓ Risco do extremo foco externo <ul style="list-style-type: none"> • Não cumprir as entregas prometidas 	Foco interno <ul style="list-style-type: none"> ✓ Mede componentes ✓ Entrega consistente, mas parcialmente eficaz ✓ Redução de custo pela consolidação ✓ Especialistas técnicos Foco externo <ul style="list-style-type: none"> ✓ Mede resultado ✓ Baixa consistência na entrega ✓ Orçamento da Unidade de Negócio ✓ Generalistas 	<ul style="list-style-type: none"> ✓ Compreender que serviços são usados pelo negócio e por que ✓ Compreender como os componentes suportam os serviços ✓ Indicadores correlacionados Negócio/TI ✓ Serviços padrão ✓ Retorno sobre investimento

Fonte: (ILLUMNA FOUNDATION CERTIFICATE, 2007)

Quadro 2 Equilíbrio Estabilidade x Responsividade

Conflito	Características	Chaves da resolução
<ul style="list-style-type: none"> ✓ Visão estabilidade (entrega consistente, serviço disponível) ✓ Visão responsividade (atender novas necessidades do negócio rapidamente) ✓ Risco da extrema estabilidade <ul style="list-style-type: none"> • Ignorar novas necessidades de negócio ✓ Risco da extrema responsividade <ul style="list-style-type: none"> • Gastos excessivos com mudanças e crises associadas 	Foco estabilidade <ul style="list-style-type: none"> ✓ Foco tecnológico e refinamento dos padrões ✓ Unidades de negócio assumem novos sistemas ✓ Soluções baseadas na tecnologia já utilizada Foco responsividade <ul style="list-style-type: none"> ✓ Foco no resultado sem preocupação com meios ✓ Foco em novos projetos e não na operação ✓ Compra de novas tecnologias para novas demandas ✓ Faltam padrões 	<ul style="list-style-type: none"> ✓ Investir em tecnologias adaptáveis, abertas ✓ Processo de GNS forte ✓ Acionar o processo de Gerenciamento de Mudanças no estágio inicial do ciclo de vida do serviço ✓ Evitar acordos informais entre TI e negócio

Fonte: (ILLUMNA FOUNDATION CERTIFICATE, 2007)

Quadro 3 Equilíbrio Qualidade do Serviço x Custo do Serviço

Conflito	Características	Chaves da resolução
<ul style="list-style-type: none"> ✓ Visão qualidade (Performance, Disponibilidade, Recuperação) 	Foco qualidade <ul style="list-style-type: none"> ✓ Entrega do pedido independente do investimento 	<ul style="list-style-type: none"> ✓ Conhecer do serviço as situações em que aumento de qualidade não representa o mesmo

<ul style="list-style-type: none"> ✓ Visão custo (Otimização de recursos e custos) ✓ Risco do extremo foco qualidade <ul style="list-style-type: none"> • Gastos acima do necessário ✓ Risco do extremo foco custo <ul style="list-style-type: none"> • Perda de qualidade 	<ul style="list-style-type: none"> ✓ Aumento progressivo de gastos ✓ Entrega mais do que necessário ✓ Falta de comunicação dos custos de TI <p>Foco custo</p> <ul style="list-style-type: none"> ✓ Atingir o orçamento e reduzir custos ✓ Limitação da qualidade pelo orçamento ✓ Resultado de TI somente baseado no orçamento 	<p>aumento em custo</p> <ul style="list-style-type: none"> ✓ Ex. Disponibilidade (55 a 75%) ✓ Conhecer do serviço as situações em que pequeno aumento de qualidade representa um aumento muito grande de custo ✓ Ex. Disponibilidade (96 a 99%) ✓ Gerenciamento Financeiro de TI com GNS
---	--	--

Fonte : (ILLUMNA FOUNDATION CERTIFICATE, 2007)

Quadro 4 Equilíbrio Reativo x Proativo

Conflito	Características	Chaves da resolução
<ul style="list-style-type: none"> ✓ Visão reativa (atua quando acontece uma demanda ou falha) ✓ Visão proativa (sempre procura melhoria na situação atual) ✓ Risco do extremo foco reativo <ul style="list-style-type: none"> • Não é efetivo no suporte à estratégia do negócio ✓ Risco do extremo foco proativo <ul style="list-style-type: none"> • Excesso de mudanças e gastos 	<p>Foco reativo</p> <ul style="list-style-type: none"> ✓ Responde a necessidades de negócio e incidentes depois que acontecem ✓ Reincidência, <i>Turn-over</i> alto ✓ Projetos feitos como se fosse a primeira vez ✓ Atua na crise ✓ Mudanças não registradas <p>Foco proativo</p> <ul style="list-style-type: none"> ✓ Antecipa necessidades de negócio e evita incidentes ✓ Gastos desnecessários ✓ TI tende a pensar que sabe mais do que o negócio sobre as necessidades do mesmo 	<ul style="list-style-type: none"> ✓ Integração do Gerenciamento de Incidentes com Problemas ✓ Habilidade de priorização entre falhas técnicas e demandas de negócio ✓ Bom gerenciamento de configuração e ativos ✓ GNS na Operação de Serviço

Fonte : (ILLUMNA FOUNDATION CERTIFICATE, 2007)

2.2.3 Gerenciamento de Incidentes

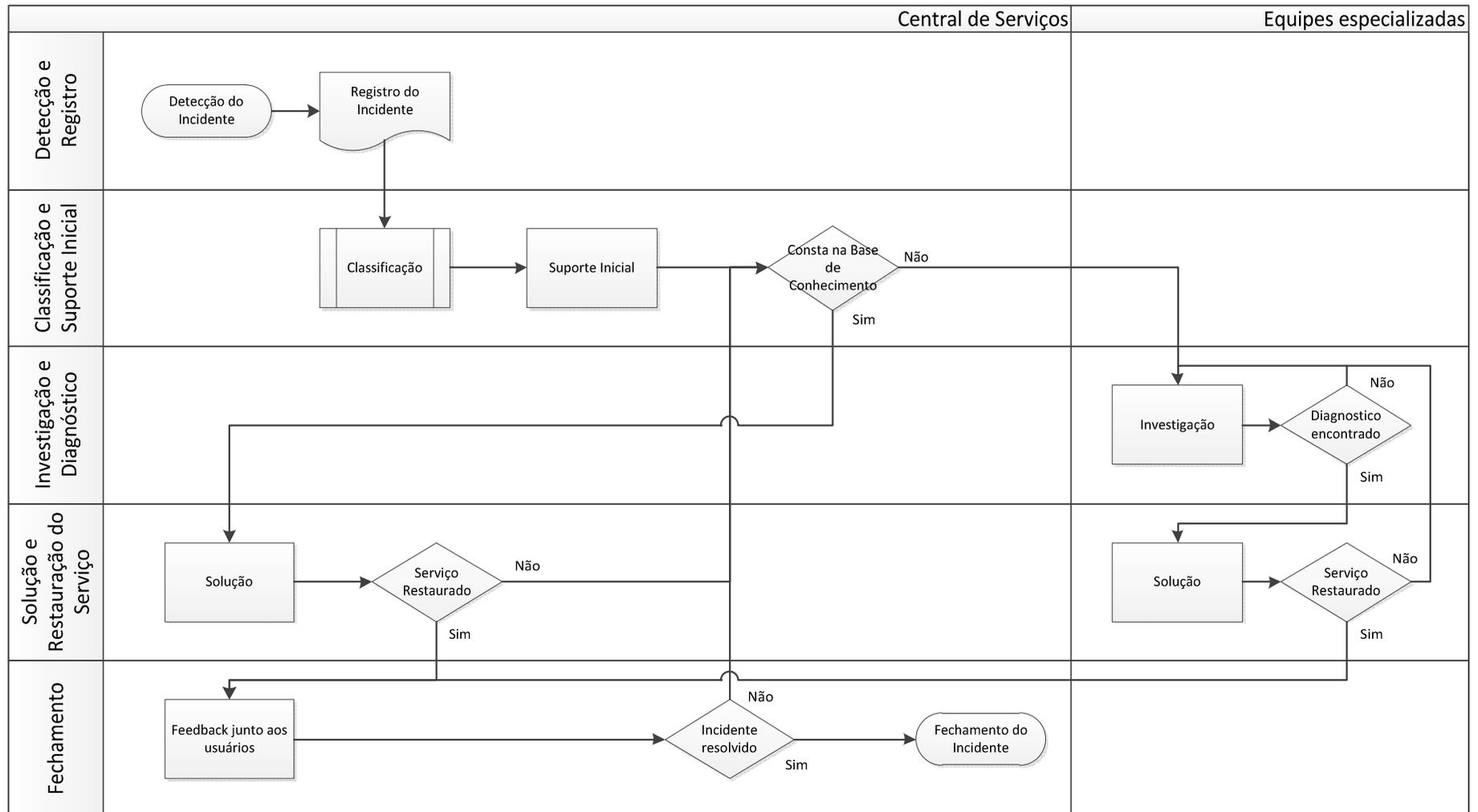
Os objetivos, conceitos e características básicas do Gerenciamento de Incidentes, incluindo incidente, escalções e priorização continuam os mesmos na ITILv3 se comparado com a segunda versão. A principal evolução é a descrição mais clara e direta do Tratamento de Incidentes (*Incident Handling*), a descrição de

procedimentos executados pela Central de Serviços e demais equipes de suporte, sendo as principais (ILLUMNA FOUNDATION CERTIFICATE, 2007):

- Detecção e registro do incidente, onde a Central de Serviços recebe ou percebe o comportamento anormal ou potencial comportamento anormal no ambiente e o registra.
- Classificação e suporte inicial por parte da Central de Serviços com base no conhecimento do ambiente e Base de Conhecimento a partir da categorização do incidente e, encaminhamento para equipes de suporte especializadas quando for este o caso.
- Investigação e diagnóstico por parte das equipes de suporte especializadas quando não se tem uma Base de Conhecimento sobre o incidente.
- Solução e restauração do serviço por parte da Central de Serviços quando uma Base de Conhecimento compatível estiver disponível ou pelas equipes de suporte especializadas quando a investigação e diagnóstico forem necessários, inserindo as informações sobre a solução na Base de Conhecimento.
- Fechamento do incidente por parte da Central de serviços, efetuando *feedback* junto aos usuários, confirmando a solução e restauração do serviço.

A Figura 3 demonstra este fluxo de trabalho de forma gráfica e resumida.

Figura 3 Fluxo de tratamento de incidentes



Fonte: Autoria própria, 2011.

Além do Tratamento de Incidentes, o Ciclo de Vida do Incidente também é definido: novo, aceito, assinalado, em progresso, resolvido e fechado. Também passamos a ter Modelos de Incidentes (*Incident Models*) que nada mais são do que rotinas, ou scripts, contendo os passos e procedimentos para tratar incidente com ordenação, responsabilidades, regras, escalação, etc, principalmente para incidentes relacionados à segurança. (ILLUMNA FOUNDATION CERTIFICATE, 2007)

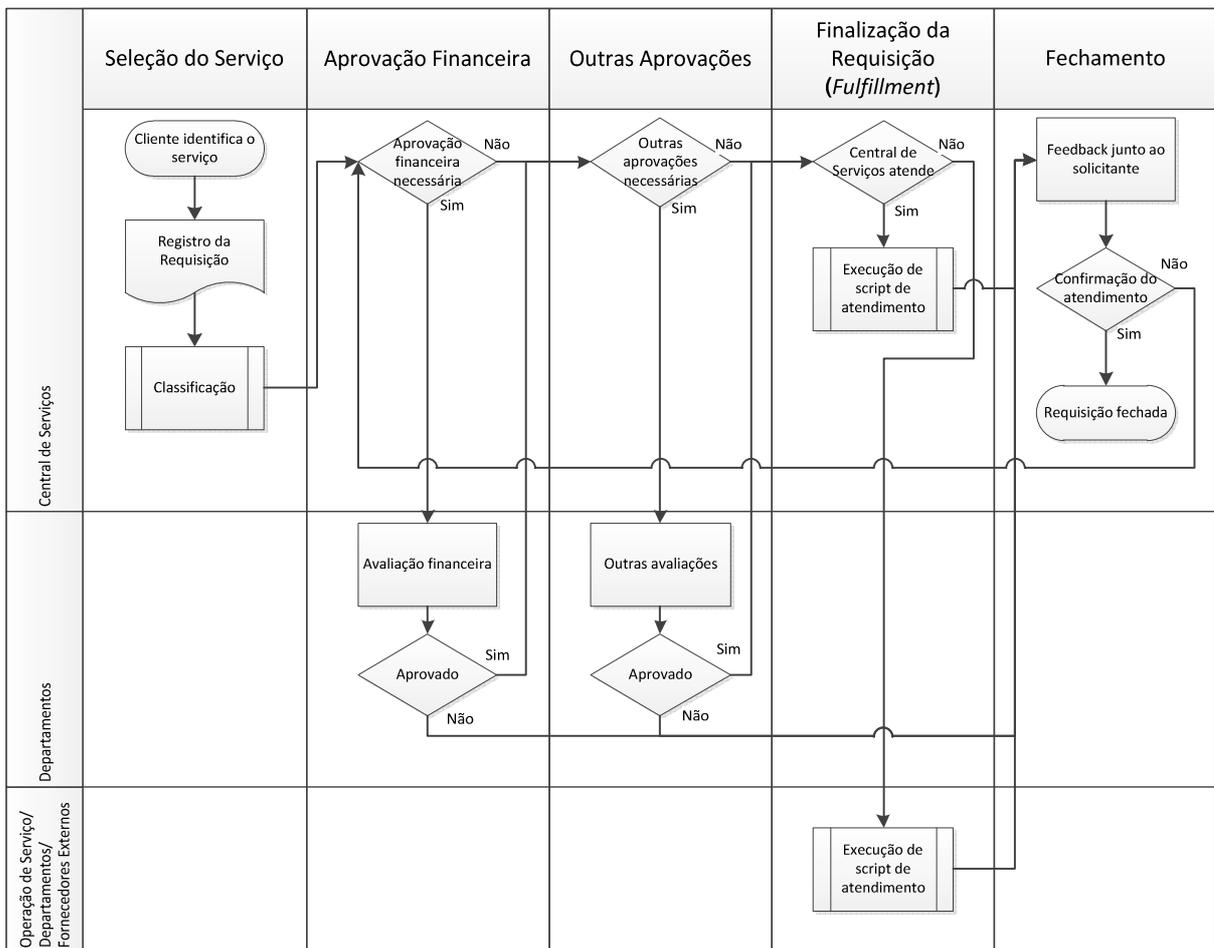
2.2.4 Cumprimento de Requisição

Uma parte importante que afeta diretamente o *Service Desk* é o tratamento de requisições, pois é por ele que a requisição é registrada caso não exista um sistema automático para tal e o “dono” da requisição sempre é o *Service Desk*. Tão grande a importância e impacto das requisições que na ITILv3 foi criado um processo apenas para seu tratamento.

O processo de Cumprimento de Requisição trata as Requisições de Serviço dos usuários durante seu ciclo de vida, fornecendo um canal para solicitações, *Service Desk* ou ferramenta automatizada, e receber serviços padronizados para aqueles que existirem um processo pré-definido de aprovação e qualificação, disponibilizando sempre informações sobre serviços disponíveis e a forma de obtê-los e conduzir reclamações e sugestões. Normalmente não é necessário adicionar papéis ou postos de trabalho, mas se o volume de requisições e/ou sua criticidade sejam altos, é recomendável ter membros da equipe de Gerenciamento de Incidentes dedicados para seu tratamento.

Segundo Illumna Foundation Certificate, (2007) uma requisição de serviço “é uma descrição genérica de vários tipos de demanda colocados pelos usuários para o departamento de TI”. Muitas delas são pequenas mudanças de custos e riscos baixos, que ocorrem frequentemente (Mudanças Padrão). É recomendado que fosse criados Modelos de Requisições para definir o fluxo padrão, com os passos necessários para atender a requisição, os envolvidos, metas, prazos de atendimento, regras de escalação, etc. O fluxo de tratamento de requisições difere significativamente do fluxo de tratamento de incidentes como podemos visualizar na Figura 4, porém sua categorização continua de suma importância.

Figura 4 Fluxo de tratamento de requisições



Fonte: Autoria própria, 2011.

2.2.5 Gerenciamento de Eventos

Este é o processo responsável pelo monitoramento de eventos que ocorrem na infraestrutura de TI para garantir sua operação normal, interpretando os eventos para detectar e escalar condições de exceção, que podem impactar no ambiente de forma negativa. Introduzido na ITILv3, pois a todo momento estão ocorrendo eventos, mas nem todos geram, ou podem gerar incidentes, bem como, mesmo não sendo incidente um evento pode necessitar ser tratado por se tratar de uma falha ou aviso de comportamento anormal de uma tecnologia. (ILLUMNA FOUNDATION CERTIFICATE, 2007)

O *Service Desk* é envolvido neste processo apenas para fins de interação com o usuário como exemplo de notificar sobre a liberação de um processamento.

2.2.5.1 Evento

Um evento é definido em Illumna Foundation Certificate, (2007) como “qualquer ocorrência detectável ou discernível que tenha significância para o gerenciamento da infraestrutura de TI ou entrega dos serviços de TI quanto ao impacto que um desvio possa causar no serviço”. Basicamente são acontecimentos relacionados à infraestrutura e/ou serviço que são monitorados e medidos, normalmente por ferramentas de monitoração, que, dependendo do contexto, pode impactar na operação normal do serviço para o negócio. A monitoração pode ser feita de duas formas separadamente ou combinadas: Ativa, onde a ferramenta de monitoria busca as informações (*polling*) ou Passiva, em agentes que enviam notificações sobre os ICs.

2.2.6 Gerenciamento de Problemas

A terceira versão da ITIL apenas incrementa o Gerenciamento de Problemas com o conceito de Base de Dados de Erros Conhecidos (BDEC ou, em inglês, *Known Error Database – KEDB*), que consiste em um repositório de informações sobre erros conhecidos, alimentada pelo processo de Gerenciamento de Problemas e utilizado por ele e pelo Gerenciamento de Incidentes e faz parte do Sistema de Gerenciamento do Conhecimento do Serviço. Outra alteração é que o registro do erro conhecido deve ocorrer quando for útil fazê-lo, na ITILv2 o registro só ocorre após a identificação da “causa-raiz” e uma solução de contorno. (ILLUMNA FOUNDATION CERTIFICATE, 2007)

2.2.7 Gerenciamento de Acesso

Outro processo operacional adicionado na ITILv3 é o Gerenciamento de Acesso, que visa garantir aos usuários autorizados o direito de usar um serviços, ao mesmo tempos que previne o acesso a usuários não autorizados, executar as políticas e ações definidas no Gerenciamento de Segurança e no Gerenciamento de Disponibilidade e garantir que a organização seja capaz de gerenciar a

confidencialidade, disponibilidade e integridade dos seus dados e ativos de propriedade intelectual. (ILLUMNA FOUNDATION CERTIFICATE, 2007)

As principais atividades do Gerenciamento de Acesso são:

- Requisição de acesso – requisição formal de criação/alteração/remoção de acesso para usuários. Pode ser via RFC ou outro mecanismo reconhecido pela empresa.
- Verificação – da requisição de acesso em duas visões: identidade do usuário e rela necessidade de acesso ao serviço.
- Prover Direitos – execução das políticas de acesso quando solicitado e verificado (o Gerenciamento de Acesso não decide as políticas de acesso).
- Monitoração do Status das Identidades – acompanhamento dos papéis das pessoas na empresa para confirmar as necessidades de acesso, sua ampliação ou revogação.
- Registro de acompanhamento de acesso – verificação constante se os direitos de acesso dos usuários estão sendo usados adequadamente.
- Removendo ou restringindo direitos – o Gerenciamento de Acesso é responsável por remover ou restringir o direito dos usuários segundo mude seu perfil de acesso.

Este processo também afeta diretamente o *Service Desk*, pois é ele que recebe as requisições de acesso, valida suas autorizações identificando a identidade do usuário, encaminhando a requisição para a equipe apropriada e detecta e reporta incidentes relacionados com acesso, sempre mantendo comunicação com o usuário.

2.2.8 Funções e papéis na Operação do Serviço

2.2.8.1 Central de Serviços (*Service Desk*)

Além das responsabilidades herdadas da ITILv2 a Central de Serviços, como vimos nos processos definidos e redefinidos na ITILv3, agregou várias novas atribuições e responsabilidades que impactam diretamente em seu dia-a-dia e, em alguns casos, até mesmo sua organização. Porém muitos destes processos definiram atividades que, em sua maioria, já passam pela Central de Serviços, mas

não eram tratados de forma tão eficiente e eficaz, pois acabam baseando-se no *feeling* das pessoas envolvidas. (RHINO CONSULTING, 2010)

2.2.8.2 Gerenciamento Técnico (*Technical Management*)

Com o objetivo de apoiar o planejamento, implementação e manutenção de uma infraestrutura de IT estável e capaz de sustentar o negócio através de uma topologia resiliente, bem arquitetada e de custo adequado utilizando habilidades técnicas eficazes, o Gerenciamento Técnico é introduzido oficialmente na ITILv3. Composto por pessoas com expertise técnica sobre infraestrutura de TI que também fazem parte dos grupos de Operação de TI e pode englobar mais de um departamento e vários times, estas pessoas são organizadas em grupos de Especialistas por Tecnologia e exercem dois papéis (ILLUMNA FOUNDATION CERTIFICATE, 2007):

- Guardar Conhecimento – são as pessoas que possuem o conhecimento técnico e expertise relacionada ao gerenciamento da infraestrutura de TI, garantindo a identificação e desenvolvimento deste conhecimento para o desenho, teste e gerenciamento dos serviços de TI, sempre em parceria com o Gerenciamento de Aplicações.
- Prover Recursos – para o Ciclo de Vida do GSTI, garantindo que hajam recursos capacitados para o desenho, construção, operação e melhoria da entrega e suporte dos serviços de TI.

2.2.8.3 Gerenciamento de Aplicações (*Application management*)

Assim como o Gerenciamento Técnico, o Gerenciamento de Aplicações possui os mesmos objetivos, papéis e sobreposições organizacionais, porém com uma visão de aplicações e aplicativos para o negócio, apoiando os processos de negócio e ajudando na identificação de requerimentos funcionais e de gerenciamento, garantindo recursos tecnicamente habilitados para tal. Seus componentes são organizados em grupos de Especialistas por Aplicação Funcional. (ILLUMNA FOUNDATION CERTIFICATE, 2007)

2.2.8.4 Gerenciamento da Operação de TI (*IT Operations Management*)

Organizacionalmente sobreposta ao gerenciamento Técnico e de Aplicações, o Gerenciamento de Operação de TI varia de acordo com a maturidade da organização de TI, com os objetivos de manter o “*status quo*” da operação, visando a estabilidade aos processos de negócio, avaliando constantemente oportunidades de melhoria para ganhar performance e reduzir custos, sendo que a capacidade técnica para diagnóstico e resolução de falhas na operação de TI deve ser adequadamente utilizada (ILLUMNA FOUNDATION CERTIFICATE, 2007). Em muitas empresas a operação de TI acaba sendo sobrecarregada com demandas extras e desvios de funções, impactando diretamente em sua eficiência e eficácia, fazendo com que a percepção de toda a TI seja abalada.

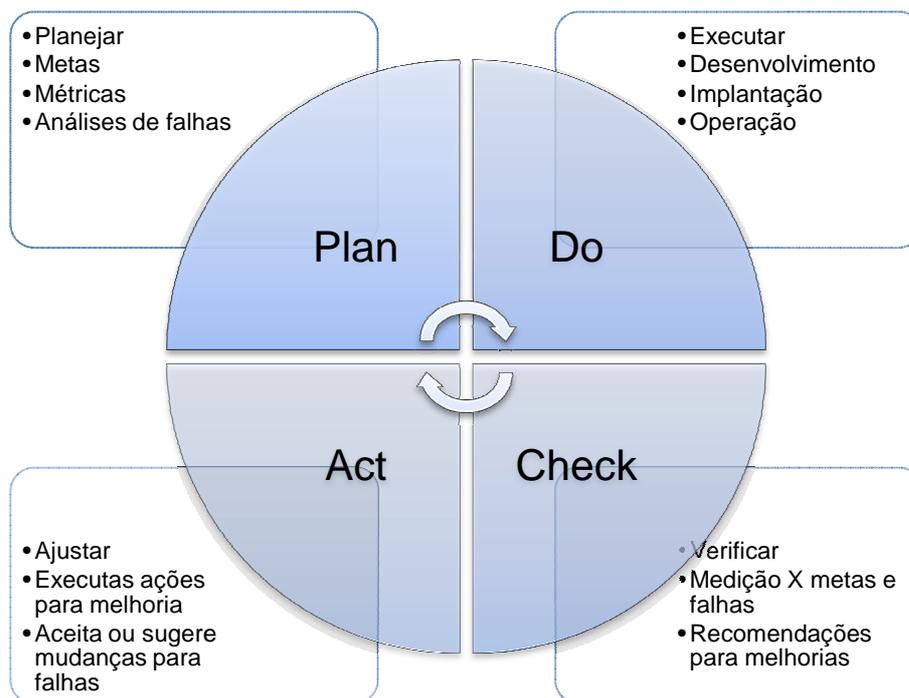
Também exerce papel duplo:

- Cuidar da Operação – executando as atividades padrão definidas, buscando estabilidade.
- Adaptação às Mudanças – diante de novos requerimentos e demandas de negócio com o mínimo de risco à operação.

2.2.8.5 Melhoria de Serviço Continuada (*Continual Service Improvement*)

Responsável por alinhar e realinhar os Serviços de TI às mudanças nas necessidades do negócio, identificando e implementando melhorias, revisar, analisar e fazer recomendações de oportunidades de melhoria em cada fase do ciclo de vida dos processos, identificar e implementar atividades para melhorar a qualidade do serviço de TI e dos processos de ITSM, reduzindo custos sem sacrificar a satisfação do cliente. Esta pessoa, ou equipe, trabalha muito próximo ao Gerente de Serviço para monitorar e auxiliar nas decisões e métricas para o negócio, principalmente quando sendo uma empresa prestadora de serviços de TI. É onde o Ciclo de Deming (Figura 5) está mais presente dentro da ITILv3.

Figura 5 Modelo PDCA aplicado à CSI



Fonte: Autoria própria, 2011.

2.3 Considerações finais

Neste capítulo foram apresentados os conceitos básicos de ITIL mais relevantes para o presente trabalho, a fim de esclarecer as relações e impactos entre os processos e funções descritos. No próximo capítulo ter-se-ão embasamento teórico sobre as técnicas e ferramentas utilizadas para tratar o problema de pesquisa, Inteligência Artificial, Inferência, Árvores de Decisão, Redes Neurais, para codificação dos dados e o Weka para os experimentos.

3 APLICAÇÃO DE INTELIGÊNCIA ARTIFICIAL EM *DATA MINING*

Grandes centros de informações em um determinado ambiente tornam-se de grande valia se usado de forma a identificar padrões, exceções, tendências e correlações. Para facilitar esta tarefa podemos utilizar a área de computação aplicada, também conhecida como Inteligência Artificial. Tornando possível detectar tendências em um ambiente que recebe a quantia de oito mil chamados por mês, como um dos ambientes estudados no presente trabalho podendo trazer benefícios como: economia, treinamento direcionado e etc.

Data Mining é uma técnica que permite identificar padrões e exceções utilizando-se de conceitos de IA e outras do conhecimento como Banco de dados, Estatística, Matemática e Sistemas distribuídos (RUSSELL e NORVIG, 2004). Uma definição sucinta encontrada em (SANTOS, 2005a), define *Data Mining* como: uso de algoritmos para extrair conhecimento de dados. Estes algoritmos são usados para:

- Classificação: aprendizado de uma função que mapeia um dado em uma de várias classes conhecidas.
- Regressão (predição): aprendizado de uma função que mapeia um dado em um valor real.
- Agrupamento (*clustering*): identificação de grupos de dados com características semelhantes entre si e diferente entre os grupos.
- Sumarização: descrição do que caracteriza um conjunto de dados (ex. conjunto de regras).
- Detecção de desvios ou *outliers*: identificação de dados que deveriam seguir um padrão, mas não o fazem.
- Análise e visualização dos resultados.

Neste sentido o *Data Mining* faz parte de um processo chamado KDD (*Knowledge Discovery in Databases*) que tem como objetivo de descobrir conhecimentos úteis previamente desconhecidos a partir de grandes bancos de dados.

Também são processos de um KDD:

- Seleção: escolha de dados para processamento.

- Pré-Processamento: enriquecimento dos dados, desnormalização de bancos de dados, etc.
- Transformações: filtragem, normalizações, reprojções, adequações a algoritmos;
- Interpretação: visualização, validação, etc.

Além do ambiente proposto neste trabalho o *Data Mining* pode ser encontrado em diversas aplicações incluindo: Economia; Detecção de fraudes; Comércio eletrônico e *marketing*; WWW; Bioinformática e medicina; Ciências sociais; Ciências da Terra.

Na presente pesquisa foram aplicadas algumas técnicas e algoritmos de inferência e classificação de dados, da área de Inteligência Artificial, que serão descritas nas próximas seções. Apresenta-se também um ambiente para análise de conhecimento denominado Weka, o qual foi utilizado na implementação deste trabalho.

3.1 Modelos de Aprendizado

Alguns algoritmos de Inteligência Artificial (RUSSELL e NORVIG, 2004) auxiliam a implementar técnicas de Data Mining, tais como por exemplo:

- Árvores de decisão – toma como entrada atributos e retorna uma decisão, valor de saída previsto segundo entrada. Os atributos de entrada podem ser discretos ou contínuos. Quando da aprendizagem for à função de valores discretos, é chamada de aprendizagem de classificação, quando de uma função contínua, é chamada de regressão. Seu algoritmo básico é definido como:
 - a) Para criar cada nó da árvore, decidimos qual atributo deve ser testado.
 - i. Escolhemos o atributo com maior ganho de informação;
 - ii. Ganho de informação relacionado com entropia;
 - b) Dividimos a árvore usando o atributo escolhido.
 - c) Executamos recursivamente.

d) Paramos quando não pudermos mais decidir (ex. quando a entropia for igual à zero).

- Descoberta de regras de associação - encontrar associações de itens ou conjuntos (*item sets*) que sejam: Significativos e Frequentes
- Vizinhos mais próximos – Método de classificação supervisionada que usa os valores das amostras de classes conhecidas;
- Agrupamento – O objetivo é achar grupos naturais nos dados. Dados no mesmo grupo são considerados similares, Dados de um grupo são considerados diferentes de dados em outro grupo.
- Mapas auto organizáveis: Usadas para mapeamento de vetores. Vetores (dados) de entrada são mapeados nas unidades ou neurônios da rede.
- Visualização – Técnicas para diferentes tipos de dados, diferentes dimensões, permitem ter uma ideia da distribuição dos dados, permitem ver um resultado de classificação/mineração.
- Listas de Decisão – As listas de decisão lembram árvores de decisão, mas sua estrutura global é mais simples, elas são expressões lógicas de formas restritas, isto é, consiste em uma série de testes, sendo que cada um é uma conjunção de literais, tornando os testes individuais mais complexos se comparados às árvores de decisão. (RUSSELL e NORVIG, 2004)

3.1.1 Máxima probabilidade: modelos discretos

Uma tarefa de aprendizagem de parâmetros envolve a descoberta de valores numéricos para um modelo de probabilidade cuja estrutura é fixa, isto é, os dados são discretos. Quando cada ponto de dados pode ser visualizado em todas as variáveis no modelo de probabilidade, dizemos que os dados são completos, o que simplifica bastante o problema de aprender um modelo complexo. (RUSSELL e NORVIG, 2004)

Matematicamente falando, aproveitando a máxima probabilidade, obtemos os valores utilizando-se de probabilidade logarítmica. Onde os parâmetros θ , hipótese h_{θ} , dados de aprendizagem N , teremos uma probabilidade logarítmica básica na Equação 1.

Equação 1 Probabilidade logarítmica básica

$$\sum_{j=1}^N \log P(d_j | h_\theta)$$

Fonte: (RUSSELL e NORVIG, 2004)

3.2 Árvore de Decisão

Sistemas especialistas utilizando árvores de decisão para o aprendizado realizam uma estratégia de divisão e conquista onde um problema complexo é composto em subproblemas mais simples. Por trabalharem em uma estrutura de árvore são necessários algoritmos recursivos, e estes são aplicados a cada subárvore ou também chamado subproblema.

Uma árvore de decisão pode crescer até todo nó ser puro (árvore máxima), quando ela terá 100% de precisão nos dados de treinamento. A árvore máxima é o resultado de superadaptação (*over fitting*), pois ela se adapta à variação sistemática do resultado (sinal) e da variação aleatória (ruído). Com isso, ela não generaliza bem os novos dados, os quais normalmente contêm muito ruído. Ao contrário, uma árvore pequena com somente poucos ramos pode subaproveitar os dados (*under fitting*) e, conseqüentemente, pode falhar na adaptação ao sinal.

Isto resulta numa generalização pobre. Para evitar estes problemas, é necessário executar a poda da árvore, que pode ser do tipo descendente ou ascendente. Na poda descendente (*top down pruning*), ou pré-poda (*pre pruning*), podem ser utilizados os seguintes critérios de parada: limite na profundidade da árvore, limite na quantidade de fragmentação (por exemplo, não dividir um nó se o número de casos ficarem abaixo de determinado limite) ou significância estatística (quando o teste chi-quadrado é utilizado como critério de quebra). Na poda ascendente (*bottom-up pruning*), também chamada de pós-poda (*post pruning*), uma grande árvore é gerada e então os ramos são cortados de maneira reversa usando um critério de seleção de modelo. A pré-poda normalmente é mais rápida, mas é consideravelmente menos eficiente que a pós-poda. (GAMA, 2002)

Outra técnica que reduz a superadaptação é a validação cruzada (*cross-validation*), onde a ideia básica é estimar até que ponto cada cálculo irá prever

dados não vistos no treinamento. Isto é feito separando-se uma fração dos dados conhecidos e usando-os para testar o desempenho de previsão de uma hipótese induzida a partir dos dados restantes. Uma validação cruzada de k vias (*folds*) significa que serão executados k experimentos separando em cada um uma fração de $1/k$ dos dados para testes diferentes em cada experimento e calcular a média dos resultados. Esta técnica pode ser utilizada com qualquer algoritmo de aprendizagem. (RUSSELL e NORVIG, 2004)

3.2.1 Algoritmos clássicos

Abaixo seguem alguns algoritmos utilizados para resolução de árvores de decisão (RUSSELL e NORVIG, 2004):

- ID3 – um dos primeiros algoritmos de árvore de decisão binária, tendo sua elaboração baseada em sistemas de inferência e em conceitos de sistemas de aprendizagem. Logo após foram elaborados diversos algoritmos, sendo o mais conhecido o C4.5.
- C4.5 – Utilizado para soluções que muitas vezes consistem em simplesmente em fornecer a classificação de um caso que lhes é apresentado. Capaz de aprender, olhando para um conjunto de casos, como eles são classificados e a partir daí fazer uma predição para novos casos. O C4.5, também binário, age como um especialista, classificando os casos desconhecidos, ele também possui um sistema pelo qual o usuário pode construir novos modelos e estudar o uso do programa nestes casos.
- J48 – Trata-se de um procedimento de descoberta de regras, binárias ou não, no qual se apresenta ao algoritmo dados com respostas conhecidas para que através destes sejam geradas regras classificatórias do tipo “se, então”, que indicam a classe mais provável para cada registro de dados, o J48 foi escolhido para este trabalho por ser o mais indicado para o problema, pois aparentemente é o que melhor está adaptado para trabalhar com árvores de decisão multivalorada. (MURALIKRISHNA, DAL LAGO, *et al.*, 2005)

3.3 Redes Neurais

Composto por 100 bilhões de neurônios o cérebro humano é considerado o processador mais complexo existente. Responsável por coordenar todos os movimentos do organismo, o cérebro, é formado por pequenas células denominadas neurônios interconectados formando uma grande rede chamada de Rede Neural. As conexões entre os neurônios, chamados de sinapses, transmitem estímulos através de diferentes concentrações de Na⁺ (Sódio) e K⁺ (Potássio), e o resultado disto pode ser estendido por todo o corpo humano. Esta grande rede proporciona uma fabulosa capacidade de processamento e armazenamento de informação.

Os principais componentes dos neurônios são (CARVALHO, 2003):

- I. Os dendritos, que tem por função, receber os estímulos transmitidos pelos outros neurônios;
- II. O corpo de neurônio, também chamado de *somma*, que é responsável por coletar e combinar informações vindas de outros neurônios;
- III. E finalmente o axônio, que é constituído de uma fibra tubular que pode alcançar até alguns metros, e é responsável por transmitir os estímulos para outras células.

3.3.1 Breve Histórico da Rede Neural Artificial

Começa por três das mais importantes publicações iniciais, desenvolvidas por: McCulloch e Pitts (1943), Hebb (1949), e Roseblatt (1958). Estas publicações introduziram o primeiro modelo de redes neurais simulando “máquinas”, o modelo básico de rede de auto-organização, e o modelo *Perceptron* de aprendizado supervisionado, respectivamente. (THOMÉ, 2003)

3.3.2 Características das Redes Neurais

A Rede Neural é composta por diversas unidades de processamento, segundo Figura 6, conectadas através de canais de comunicação com determinado peso associado (W). Seu espaço de processamento é limitado aos seus dados

locais, identificadas pelas entradas recebidas por suas conexões (a). O resultado final, ou seja, o comportamento inteligente é o resultado das interações, de j até i , entre as unidades de processamento (N).

A operação de uma unidade de processamento, proposta por McCullock e Pitts em 1943, pode ser resumida da seguinte maneira (THOMÉ, 2003):

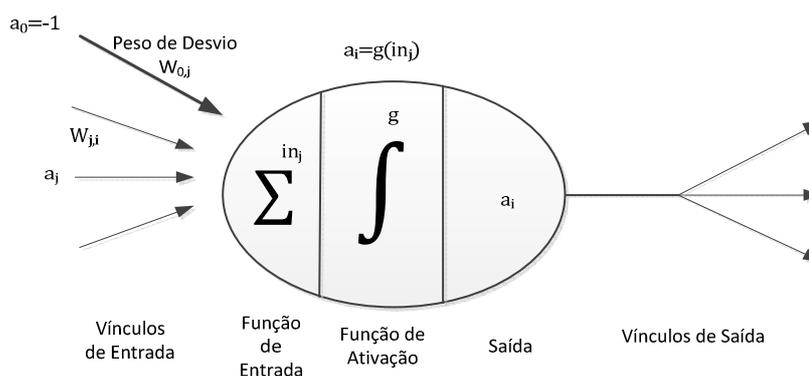
- Sinais são apresentados à entrada;
- Cada sinal é multiplicado por um número, ou peso, que indica a sua influência na saída da unidade;
- É feita a soma ponderada dos sinais que produz um nível de atividade;
- Se este nível de atividade exceder certo limite (*threshold*), definido pela Equação 2, que pode ser linear ou sigmoide, a unidade produz uma determinada resposta de saída.

Equação 2 Função de ativação

$$a_i = g(in_i) = g\left(\sum_{j=0}^N W_{j,i} a_j\right)$$

Fonte: (RUSSELL e NORVIG, 2004)

Figura 6 Modelo matemático simples para um neurônio

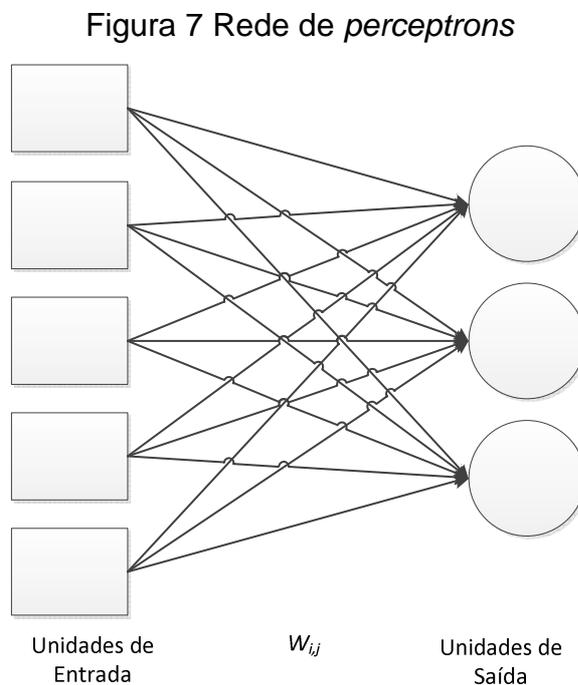


Fonte: (RUSSELL e NORVIG, 2004)

3.3.2.1 Perceptron

Uma rede neural de uma única camada, ou rede de *perceptron*, é uma rede onde todas as entradas são conectadas diretamente às saídas, segundo Figura 7.

Desta forma cada peso afeta apenas uma das saídas, pois cada unidade de saída é independente das outras.



Fonte: Adaptado de (RUSSELL e NORVIG, 2004)

Quando uma *perceptron* utiliza funções de ativação de limiar, ela é chamada de separador linear, pois tratam muito bem apenas funções linearmente separáveis, mas limita seu poder de expressão, não suportando algumas funções booleanas como, por exemplo, o XOR. (RUSSELL e NORVIG, 2004)

A ideia por trás da maioria dos algoritmos de aprendizagem de redes neurais é ajustar os pesos da rede para minimizar algum erro no conjunto de treinamento e sua medida clássica é a soma dos erros quadráticos, e aplicando cálculo de derivada em relação a cada peso, estaremos usando o declínio de gradiente para reduzir os erros quadráticos.

3.3.2.2 *Multi-layer perceptron* (MLP)

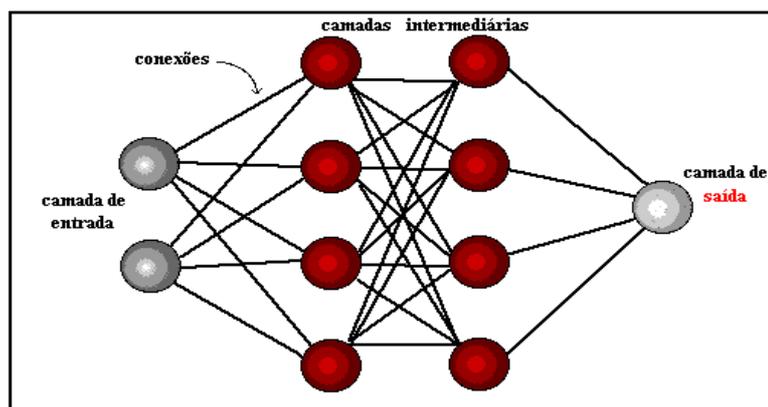
É constituída de uma rede fortemente conectada, de múltiplas camadas, de forma em que estas estão organizadas de modo que os neurônios de uma camada estimulam todos os neurônios da camada seguinte. Sendo que este estímulo só se dá entre camadas diferentes e nunca um neurônio estimula um neurônio da mesma

camada e de camadas anteriores, conforme exemplificado na Figura 8. Diferentemente de uma rede de camada única, uma MLP pode representar qualquer função, desde que tenha neurônios suficientes.

Suas características são:

- Uma camada de entrada: consiste em uma camada com os sinais de entrada (estímulo da rede). Esta camada não possui neurônios;
- Uma camada de saída: consiste em uma camada de neurônios que geram a saída da rede (resposta da rede a um estímulo);
- Camadas escondidas ou intermediárias: qualquer camada que se encontre entre a camada de entrada e saída. Não existem limites para quantidade de camadas escondidas, e também não é obrigatória à existência delas.

Figura 8 Organização das camadas em uma MLP



Fonte: (THOMÉ, 2003)

3.3.2.2.1 Treinamento

Dentre os algoritmos disponíveis para treinamento de uma MLP o mais utilizado é o proposto por Rumelhart e McClelland em 1986, conhecido como *backpropagation* (retro propagação). A arquitetura da rede deve ser previamente conhecida, e esta não é alterada durante o treinamento, onde ao longo do aprendizado os únicos parâmetros que são alterados são os pesos da rede. (NIEVOLA, 1999)

O algoritmo de *backpropagation* utiliza um algoritmo supervisionado estático, e exige em geral muitas iterações, chamadas épocas, para convergir. As principais características do *backpropagation* são:

- A entrada é apresentada e propagada para frente através da rede calculando as ativações para cada unidade de saída;
- Cada unidade de saída é comparada com o valor desejado, resultando um valor de erro;
- Existe um passo de retorno na rede onde se calculam os erros em cada unidade e são realizadas alterações nos pesos.

Algoritmo de Retro propagação pode ser assim descrito (NIEVOLA, 1999):

- I. Escolher um pequeno valor positivo para o tamanho do passo, e assinalar pesos iniciais pequenos aleatoriamente selecionados $\{w_{i,j}\}$ para todas as células.
- II. Repetir até que o algoritmo convirja, isto é, até que alterações nos pesos e no erro médio quadrático tornem-se suficientemente pequenas:
 - a. Escolher o próximo exemplo de treinamento E e sua saída correta C (a qual pode ser um vetor).
 - b. Passo de propagação avante: Fazer uma passagem da entrada para a saída através da rede para calcular as somas ponderadas, S_j , e as ativações, $u_j = f(S_j)$, para todas as células.
 - c. Passo de retro propagação: Iniciando com as saídas, fazer uma passagem de cima para baixo através das células de saída e intermediárias calculando a Equação 3.
 - d. Atualizar os pesos: $w_{ij}^* = w_{ij} + \rho \delta_i u_j$

Equação 3 Cálculo de retro propagação

$$f'(S_i) = u_i(1 - u_i) \text{ e } \delta_i = \begin{cases} (C_i - u_i)f'(S_i), & \text{se } u_i \text{ é uma unidade de saída} \\ (\sum_{m>i} w_{mi}\delta_m) f'(S_i), & \text{para outras unidades} \end{cases}$$

Fonte: (NIEVOLA, 1999)

Existem, porém, alguns problemas da retro propagação:

- Parada da rede: Se os pesos forem ajustados em valores muito grandes a ativação se torna zero ou um e os ajustes passam a ser nulos, parando a rede;

- Mínimos locais: A superfície de erro de uma rede complexa é cheia de montanhas e vales. Desta forma, a mesma pode ficar presa em um ponto de mínimo local.

Depois que a rede estiver treinada e o erro estiver em um nível satisfatório, ela poderá ser utilizada como uma ferramenta para classificação de novos dados. Para isto, a rede deverá ser utilizada apenas no modo progressivo (*feed-forward*). Ou seja, novas entradas são apresentadas à camada de entrada, são processadas nas camadas intermediárias e os resultados são apresentados na camada de saída, como no treinamento, mas sem a retro propagação do erro. A saída apresentada é o modelo dos dados, na interpretação da rede.

3.4 O Weka

O pacote *Waikato Environment for Knowledge Analysis* (Weka) é formado por um conjunto de implementações de algoritmos de diversas técnicas de Mineração de Dados. Dentre as várias funcionalidades da ferramenta (experimentações, análises automáticas, exploração de dados e fluxo de conhecimento), será utilizada a exploração dos dados para um controle mais manual e um conjunto maior de filtros. (SANTOS, 2005b)

O Weka está implementado na linguagem Java, que tem como principal característica ser portátil, desta forma pode rodar nas mais variadas plataformas e aproveitando os benefícios de uma linguagem orientada a objetos como modularidade, polimorfismo, encapsulamento, reutilização de código dentre outros, além disso, é um software de domínio público. Dentre as suas funcionalidades podemos encontrar os métodos de classificação, árvore de decisão induzida, regras de aprendizagem, redes neurais *multi-layer perceptron*, normalização e conversão (SANTOS, 2005b). No presente trabalho são utilizadas as classes de funções de classificação *J48* e *Multi-layer Perceptron*. Os dados para análise podem ser passados ao Weka de várias formas, desde acesso a banco de dados via JDBC até arquivos CSV (*Comma Separated Values*), porém a entrada padrão é um arquivo ARFF (*Attribute-Relation File Format*) que é um arquivo de texto plano composto por três partes:

- Relação: a primeira linha do arquivo deve sempre começar com a sentença *@relation* seguida por uma identificação do objeto de estudo.
- Atributos: as linhas após a relação são os atributos, cada linha deve começar com *@attribute* depois o nome do atributo, os atributos podem ser de dois tipos: numérico (onde o nome do atributo deve ser seguido pela palavra-chave *real*), ou nominal, neste caso todas as alternativas possíveis para o atributo devem estar em uma lista separada por vírgulas dentro de chaves. Preferencialmente devem-se colocar os atributos em ordem de precedência, onde o último atributo seja a classe de instância.
- Dados: depois dos atributos temos a parte de dados, que é identificada por uma linha contendo *@data*, logo na linha abaixo de *@data*, cada linha deve ser uma instância dos atributos, na mesma ordem que foi declarada na parte dos atributos, separada por vírgula.
- Podemos ter linhas de comentário, estas devem ser iniciadas com um sinal de percentagem (%). (SANTOS, 2005b)

3.5 Considerações finais

Os principais motivos para a escolha do Weka como ferramenta de experimentação são o suporte aos algoritmos escolhidos e a modularidade por ser desenvolvido em Java, permitindo chamadas diretas às classes dos algoritmos.

Com o fim deste capítulo concluímos a fundamentação teórica necessária para a compreensão do próximo capítulo que trata das características gerais dos processos e bases de dados obtidos junto às empresas, como foram escolhidos os algoritmos, a modelação e transformações necessárias dos dados e os resultados dos experimentos e suas análises.

4 DESENVOLVIMENTO

Neste capítulo são expostos os contextos referentes aos processos e características das bases de dados utilizadas nas experimentações. Indica-se a escolha dos algoritmos para análise, os tratamentos e formatação das bases de dados, bem como a execução dos experimentos. Finalmente, são apresentados os resultados obtidos e a análise a respeito dos mesmos.

4.1 Características das bases de dados

As experimentações, avaliações e métricas terão como base as informações de duas categorizações e classificações reais utilizadas em empresas diferentes e modeladas baseadas em versões diferentes da ITIL, isto é, uma tem um foco mais operacional, uma visão de tecnologia, e outra, um foco mais na percepção do usuário, cliente e negócio.

Ambas as empresas são grandes indústrias em diferentes regiões brasileiras, com características de negócio similares onde a TI é uma ferramenta, não é o foco do negócio. A quantidade de incidentes registrados é diretamente proporcional à quantidade de computadores clientes no ambiente, assim como a equipe de TI, onde os processos operacionais foram terceirizados com empresas nas quais a prestação de serviços em TI é o foco do negócio e estão capacitadas para tal.

Na primeira empresa, aproximadamente 60% maior que a segunda, a Gestão de TI modelou seus processos com base na ITILv2, pois esta revisão foi feita em 2004 e a terceira versão da ITIL ainda não fora publicada, não foram feitas revisões posteriores com base na ITILv3. Nem todos os processos ITILv2 chegaram a ser modelados, e outros foram fundidos, como, por exemplo Gerenciamento de Mudanças e Liberações, bem como as equipes e gestores sobrepuseram-se com papéis de Donos, Gerentes de Processo, equipe operacional, acabaram caindo na “armadilha” de tratar certas situações com visão equivocada e sendo praticamente todos técnicos e com foco operacional, incluindo-se os gestores, o desenho do Gerenciamento de Incidentes e conseqüentemente da categorização tende fortemente para o lado tecnológico. A classificação acabou ficando em segundo plano, tanto que ela não influencia na categorização, estando esta toda disponível, independente da classificação. Na modelagem efetuada a categorização e

classificação não são ICs e, portanto não passam pelos processos de Gerenciamento de Mudanças e Liberações, e seu fluxo de alteração é muito dinâmico. Atualmente existem 686 combinações de classificação e categorização diferentes nesta base.

Com base nesta modelagem, o *Service Desk* deve registrar o contato do usuário/solicitante, classificando-o de acordo com a visão do usuário/solicitante com relação à sua questão, isto é, se ele está com um problema e não sabe o que é e como resolver, se ele está com um problema e sabe o que é, mas não sabe como resolver, etc. Já a categorização do registro, o *Service Desk* deve fazer com base na experiência, treinamento ou *feeling* sobre o que pode ser a causa-raiz ou tecnologia mais envolvida com o registro. Estas características aumentam em muito a probabilidade de desvio nesta atividade, portanto é de extrema importância que a classificação e categorização desenhadas possuam tendência a minimizar este risco.

A segunda empresa já utiliza ITIL como base de seus processos desde 2004, e vem remodelando seu processos frequentemente, aproximadamente a cada dois anos, o que a levou, em suas últimas revisões, a introduzir alterações segundo recomendações da ITILv3. Todos os papéis de donos dos processos, embora venham da área tecnológica, são formados gestores, com uma visão de negócio, o que facilitou esta introdução das recomendações ITILv3, inclusive os gestores das funções. Os processos mais operacionais como Gerenciamento de Incidentes, Requisições, Operação, Acessos e Eventos são executados exclusivamente por equipes terceirizadas, já os Gerenciamentos de Liberação, Mudança e Acesso possuem um maior envolvimento dos gestores, até mesmo em sua execução. Assim como a primeira empresa foi optado por unir os processos de Gerenciamento de Mudanças e Liberações. A função de *Service Desk* é um serviço terceirado à parte, incluindo a gestão, estando geograficamente distante da empresa. As demais funções são compartilhadas entre equipe terceira e própria, excluindo-se o Gerenciamento de Aplicações que apenas funcionários próprios em nível de gestão se responsabilizam, acionando diversos fornecedores externos quando necessário. O Gerenciamento de configuração é bem extenso, contemplando desde hardwares e softwares, até mesmo documentações e parâmetros de processo, incluindo a classificação e classificação dos registros efetuados pelo *Service Desk*, desta forma

toda e qualquer alteração nestes deve passar pelos Processos de Gerenciamento de Mudanças e Liberações, fazendo com que passem por uma avaliação mais abrangente e cuidadosa quando uma alteração ou revisão de classificação e categorização seja necessária. Outros pontos de destaque na categorização desenhada é a quantidade de níveis, mais de 80% possui apenas dois níveis, e o impacto no SLA, pois dependendo da categorização, além das equipes solucionadoras, o tempo para atendimento, a representatividade nos KPIs e a escalação vertical podem variar. Atualmente existem 304 combinações diferentes de classificação e categorização nesta base.

Para a equipe de *Service Desk*, a classificação e categorização dos registros são mais simples se comparado com a da primeira empresa, pois a categorização está vinculada à classificação, isto é, após o registro ser classificado, como incidente, requisição, mudança padrão, etc, ficam disponíveis apenas categorizações cabíveis para esta classe de registro, reduzindo a quantidade de opções que o atendente tem a sua disposição e reduzindo a probabilidade de desvios. Embora a equipe de *Service Desk* não passe por um treinamento extensivo sobre as características da segunda empresa, é possível um bom nível de categorizações corretas, já que é baseado na percepção do usuário, o que está ocorrendo com o usuário, qual o impacto para o seu trabalho e conseqüentemente para o negócio do cliente. Uma das atribuições da equipe de Melhoria Continuada com relação ao *Service Desk* é efetuar *feedback* junto aos usuários/solicitantes de cada registro, incluindo métricas sobre o preenchimento do registro, avaliando os desvios e suas causas, neste casos recomendando planos de ações como treinamento, integração e revisão da classificação e categorização.

O Quadro 5 apresenta um resumo das características básicas que afetam ou são afetados pela Classificação e Categorização referentes às bases de dados estudadas.

Quadro 5 Características básicas das bases de dados de estudo

	Base 1	Base 2
Baseada em	ITILv2	ITILv3
Passa por processo de Gerenciamento de Mudanças	Não	Sim
Classificação afeta SLA	Sim	Sim
Categorização afeta SLA	Não	Sim
Classificação influencia equipe solucionadora	Não	Não
Categorização influencia equipe solucionadora	Sim	Sim
Classificação influencia Categorização	Não	Sim
Níveis de Categorização	3	2
Tamanho (classificações e categorizações deferentes)	686	304

Fonte: Autoria própria, 2011.

4.2 Escolha dos algoritmos

Levando em consideração as características das bases de dados, onde cada nível e sub nível possui múltiplas possibilidades, não tendo características binárias, são necessários algoritmos sem limitação a apenas funções binárias, excluindo desta forma árvores de decisão ID3 e C4.5. Redes *perceptron* de camada única também não são recomendadas já que os dados não são lineares. Árvores de decisão J48 e redes neurais *multi-layer perceptron*, segundo o que foi apresentado, são boas escolhas para avaliação. Foi visto que algoritmos genéticos e máquinas de núcleo, também chamadas de máquinas de vetor de suporte (MVS), são igualmente opções para avaliações futuras.

4.3 Configurando base de dados

Depois de selecionada a ferramenta (Weka), e recebidas as bases de dados completas exportadas dos sistemas de gerenciamento de incidentes para o formato CSV, a base de dados da primeira empresa, que será chamada de Base 1, fora normalizada para inclusão do fator Classificação (1ª nível da análise), pois este não está vinculado aos outros fatores dentro do software utilizado pelo *Service Desk* originalmente. A Base 1 também passou por uma revisão verificando as entradas com duplicidade de sentido ou grafias díspares, regularizando os desvios. Assim

ambas as bases de dados podem ser pré-processadas de uma forma equivalente para comparação. Revisados os dados crus, as tabelas de dados foram carregadas para bases de dados SQL, assim é possível uma melhor manipulação utilizando os recursos da linguagem.

A partir das bases SQL iniciou-se a formatação da base para arquivo ARFF, a primeira, levando em consideração que a ferramenta não aceita dados com espaços, portanto foram substituídos todos os espaços por *underline* (). A base de dados da segunda empresa, que será chamada de Base 2, contém caracteres especiais como “@”, “©”, “®”, etc e passou por um processo de substituição dos caracteres não suportados pelo arquivo ARFF, bem como a entrada de dados faz diferenciação entre letras maiúsculas e minúsculas (*case sensitive*), e ambas as questões foram tratadas utilizando-se de funções SQL e visões (*views*) para tal.

Durante as execuções dos primeiros experimentos, identificou-se um limitador considerável: poder computacional para treinamento da rede MLP. Assim foram direcionados esforços para encontrar uma possível solução, pois com este limitador seria inviável a aplicação de redes MLP no dia-a-dia dos processos nas empresas.

Uma solução plausível seria converter os valores das classificações e categorizações de texto (*strings*) para valores numéricos, porém deveriam ser mantidas as características discretas das bases e os relacionamentos de seus fatores. Sendo assim, optou-se por codificar os dados em produtos de números primos de fator um, baseado no Teorema Fundamental da Aritmética (vide Apêndice A). Embora não seja escopo deste trabalho, ou vantagem desta codificação é a possibilidade de identificação da categorização e classificação resultante dos algoritmos pode ser avaliada através de decodificação direta valendo-se do Teorema de Euclides para identificação de decomposição do mdc. Para tanto foi obtida uma lista com os primeiros dez mil números primos e criada uma tabela para armazená-la incluindo um índice de números inteiros positivos. (CALDWELL, 2009)

A codificação das bases para matrizes de números primos foi executada em *Transact SQL* (T-SQL) da seguinte forma:

- i. Importada lista de números primos
- ii. Base 1
 - a. Para cada nível da categorização e classificação foram identificadas as entradas

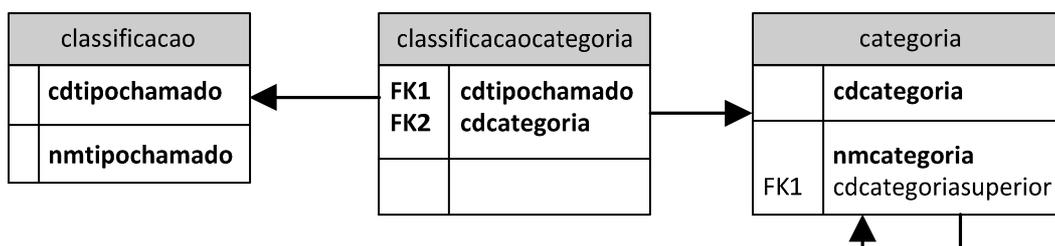
- distintas e criadas tabelas com índice de números inteiros positivos;
- b. Criada matriz de classificação e categorização com os índices utilizando visões;
 - c. Criada matriz de classificação e categorização substituindo os índices por números primos de mesmo índice utilizando visões;
- iii. Base 2
- a. Criada uma visão da Base 2 a formatar à estrutura da Base 1, como na Figura 9, pois os vínculos entre os níveis originalmente utilizavam recursividade em uma mesma tabela com indicadores numéricos como na Figura 10.
 - b. Para cada nível da categorização e classificação foram identificadas as entradas distintas e criadas tabelas com índice de números inteiros positivos;
 - c. Criada matriz de classificação e categorização com os índices utilizando visões;
 - d. Criada matriz de classificação e categorização substituindo os índices por números primos de mesmo índice utilizando visões;

Figura 9 Base 1 – Tabela de dados

categorizacao e classificacao1	
	Classificação Categoria Tipo Item

Fonte: Autoria própria, 2011.

Figura 10 Base 2 – Relacionamentos



Fonte: Autoria própria, 2011.

Com as matrizes de números primos exportadas do SQL, foram gerados os arquivos ARFF para entrada no Weka, estes estão disponíveis em APENDICE B – BASE 1 e APÊNDICE C – BASE 2.

4.4 Experimentações

Assim que os dados foram inseridos no Weka, foi iniciado o passo de pré-processamento dos dados, onde podemos aplicar filtros e totalizadores baseados em qualquer atributo ou combinação destes. Um fator impactante nos experimentos é o poder computacional necessário. Identificou-se durante os testes preliminares que o treinamento da rede neural MLP com valores do tipo texto (*string*) necessita de uma grande quantidade de memória, ao efetuar uma avaliação básica da Base 1 com todos os 4 níveis, um servidor com 16GB de memória e 8 núcleos de processamento a 2.4GHz não foi o suficiente para processar os dados, o fator limitante foi a alocação de memória para conversão de *strings* discretas para valores numéricos fracionários entre -1 e 1 utilizados pela classe do Weka para a inferência.

Em função desta dificuldade foi adotada a inferência dos dados como uma matriz de números inteiros positivos, porém neste caso, houve problemas no treinamento com relação à identificação dos valores como discretos, para o algoritmo estes números inteiros positivos eram uma função contínua, caracterizando uma regressão, portando os resultados estavam distorcidos. Outra solução adotada após pesquisas foi a utilização de produtos de fatores de números primos, pois são identificadores únicos, e não números compostos e com uma

progressão definida, a partir do Teorema Fundamental da Aritmética (vide Apêndice A), o que resolveu a dificuldade de treinamento da rede neural MLP, todavia o algoritmo de árvore de decisão J48 não conseguia tratar os valores, pois para ele, não havia mais de um atributo para relacionamento. A solução definitiva foi, em vez de utilizar o resultado de produtos de fatores de números primos, o uso apenas destes em fatores de 1 sem efetuar o cálculo, deixando o relacionamento entre os fatores a ser tratado pelos algoritmos. Assim as limitações de memória e processamento, bem como de codificação, não mais afetaram as experimentações.

O equipamento utilizado para os experimentos possui 4 núcleos de 2,7GHz e 8GB de memória, sendo que a utilização máxima durante os experimentos foi de 387MB, uma redução de mais de 97,35%.

Após o pré-processamento passamos a executar os experimentos com base em todos os atributos, aprendido de máxima probabilidade em modelos discretos, para verificar os melhores resultados tanto com os algoritmos de árvores J48 quanto *Multi-layer Perceptron*, ambos utilizando, para reduzir o ruído, *cross-validation* e a árvore com pós-poda.

Foram criados *scripts* em *PowerShell* para as iterações dos testes e formatação dos resultados para avaliação, cada script trata os experimentos utilizando chamadas para as classes Java do Weka passando as configurações e arquivo de entrada ARFF como parâmetros. Os as configurações utilizadas para cada algoritmo foram as seguintes:

- I. Árvore de decisão J48
 - I.1. Classe: `weka.classifiers.trees.J48`
 - I.2. Valores padrão
 - I.2.1. *Confidence Factor* (utilizado como margem para poda): 0.25
 - I.2.2. *Minimum Leaf* (define a quantidade mínima de folhas para uma subárvore): 2
 - I.3. *Folds* (quantidade de vias para *cross-validation*): variando de 10 em 10 desde 10 até 100 *folds*
- II. Rede neural MLP
 - II.1. Classe: `weka.classifiers.functions.MultilayerPerceptron`
 - II.2. Valores padrão

- II.2.1. *Auto Build* (conecta todos os neurônios de uma camada com todos os ad próxima automaticamente): *True*
- II.2.2. *Learning Rate* (taxa de aprendizado para cálculo da atualização dos pesos nas épocas): 0.3
- II.2.3. *Momentum* (aplicado aos pesos durante o cálculo nas épocas para minimizar problemas de mínimos locais): 0.2
- II.3. *Folds* (quantidade de vias para *cross-validation*): variando de 10 em 10 desde 10 até 100 *folds*
- II.4. *Training Time* (quantidade de épocas para o treinamento de *back-propagation*): variando de 25 em 25 desde 25 até 500 épocas
- II.5. *Hidden Layers* (Quantidade de camadas de neurônios ocultas na rede): variando de 1 em 1 desde 1 até 100 níveis

Para cada experimento fora gerado um arquivo de texto com as informações de resultado, além destes, o *script* em *PowerShell* consolidou os resultados dos experimentos agrupando por bases versus algoritmo em arquivos CSV para facilitar a compilação e análise dos resultados.

4.5 Análise dos resultados

Logo no passo de pré-processamento dos dados já foi identificado uma melhoria a ser implementada na Base 1: a vinculação da classificação aos outros três atributos, com isto têm-se uma redução significativa na quantidade de hipóteses de categorização, no pior caso de 636 para 314 e, no melhor caso, para 76, mesmo sem alterar nada na atual categorização.

O principal fator de avaliação é a coerência dos algoritmos em identificar corretamente a classificação dos dados, o valor mínimo aceitável para ambas as empresas é de 80%, segundo entrevista com os Gerentes de *Service Desk*. Os valores obtidos utilizando árvores de decisão J48 para ambas as bases de dados atendem este requisito, como pode ser visualizado na Tabela 1.

Tabela 1 Resultados de treinamento Algoritmo de Árvore de Decisão J48

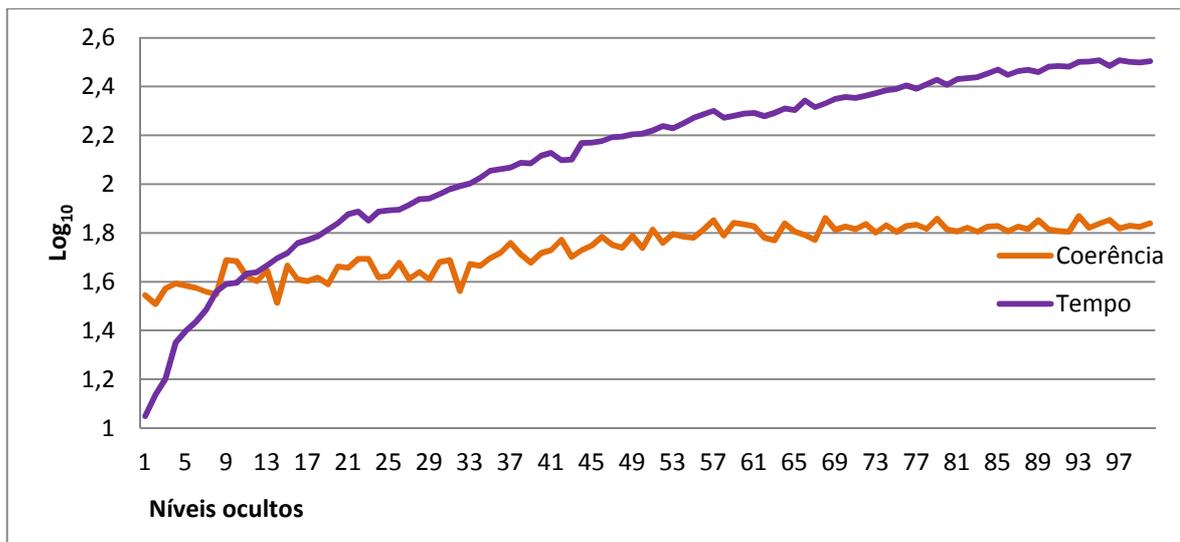
Validação Cruzada Vias	Base 1 Coerência %	Base 2 Coerência %
10	86,7925	97,6974
20	86,9497	97,6974
30	86,9497	97,6974
40	88,0503	97,3684
50	88,0503	97,3684
60	87,4214	97,6974
70	88,0503	97,3684
80	88,0503	97,3684
90	88,0503	97,3684
100	88,0503	97,3684

Fonte: Autoria própria, 2011.

Para avaliação da rede neural *multi-layer perceptron*, primeiramente é necessário identificar os parâmetros ótimos de execução do algoritmo para cada uma das bases de dados, sempre com base na relação tempo de treinamento versus coerência. Esta relação, para fins de visualização gráfica, é utilizada notação \log_{10} .

O primeiro parâmetro avaliado foi quantidade de níveis ocultos de neurônios, para a Base 1 foi identificado que 9 níveis ocultos apresentam os melhores resultados para a relação, como exibido no Gráfico 1, para a Base 2 o melhor relacionamento foi identificado como sendo no 7º nível, vide Gráfico 3. Identificada a quantidade de níveis ocultos de neurônios para cada base, foram iniciados experimentos para identifica quantas épocas de cálculo apresentam a melhor relação, 75 épocas no caso da Base 1 conforme Gráfico 2, e 25 épocas no caso da Base 2, vide Gráfico 4. O último parâmetro a ser identificado o valor de melhor relacionamento é a quantidade de vias para a validação cruzada, que se apresentou a 20 vias para Base 1, vide Gráfico 5, e 10 vias para Base 2, vide Gráfico 6.

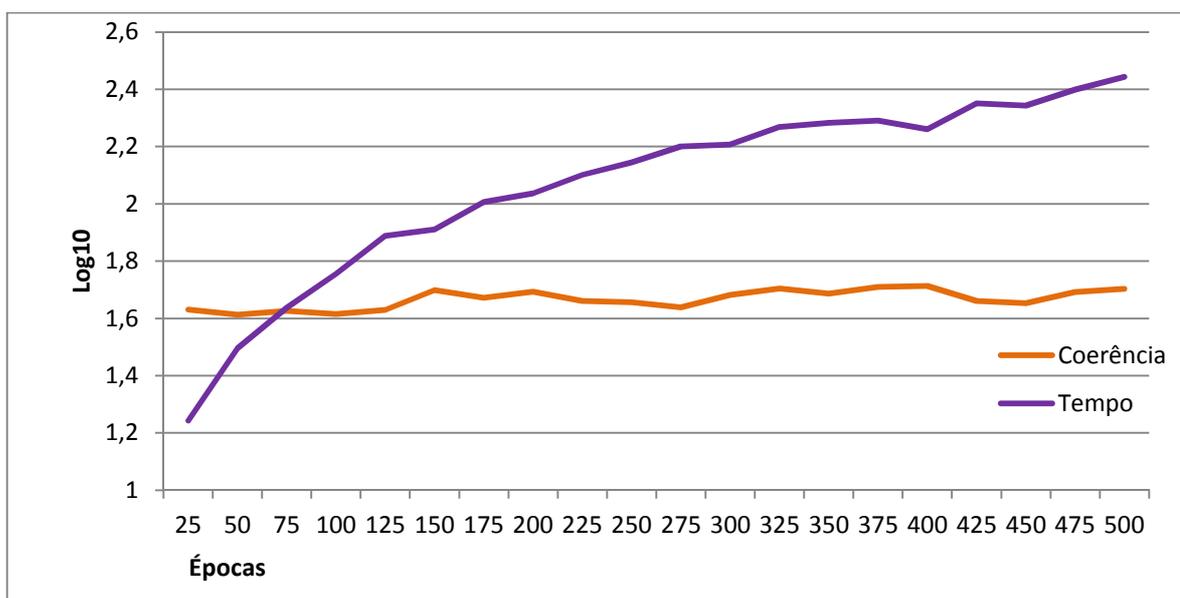
Gráfico 1 Treinamento da Base 1 – Níveis ocultos: coerência x tempo



Fonte: Autoria própria, 2011.

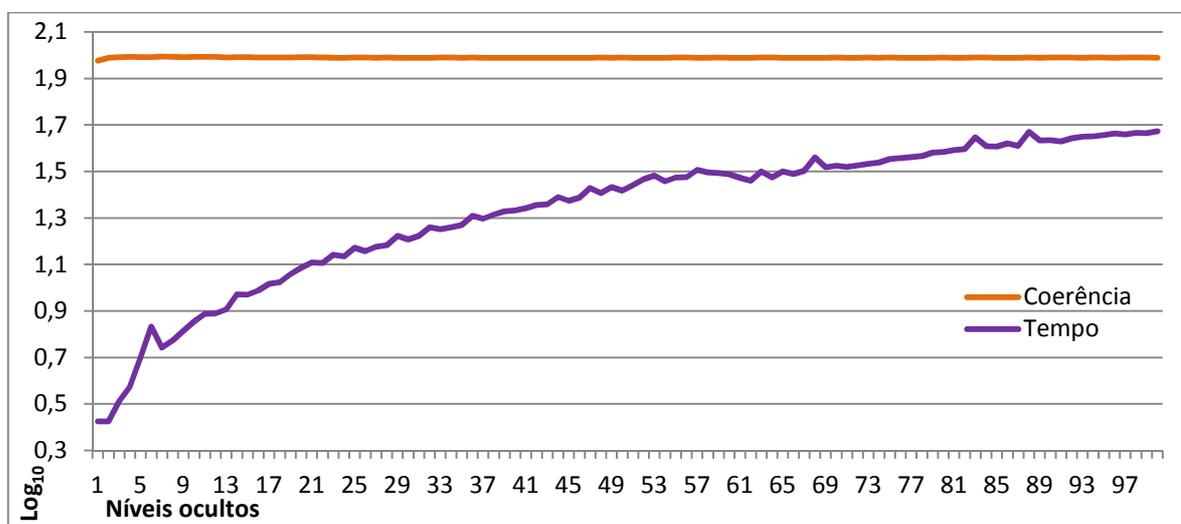
O Gráfico 1 (acima) apresenta os resultados dos experimentos executados na Base 1 utilizando rede MLP variando a quantidade de níveis ocultos de 1 a 100, demonstrando os resultados em base logarítmica da relação coerência *versus* tempo de treinamento. Neste sentido o Gráfico 2 (abaixo) demonstra os resultados da mesma relação da Base 1 no treinamento da MLP, também em base logarítmica, variando a quantidade de épocas de retro propagação de 25 a 500 em intervalos de 25 épocas.

Gráfico 2 Treinamento da Base 1 – Épocas: coerência x tempo



Fonte: Autoria própria, 2011.

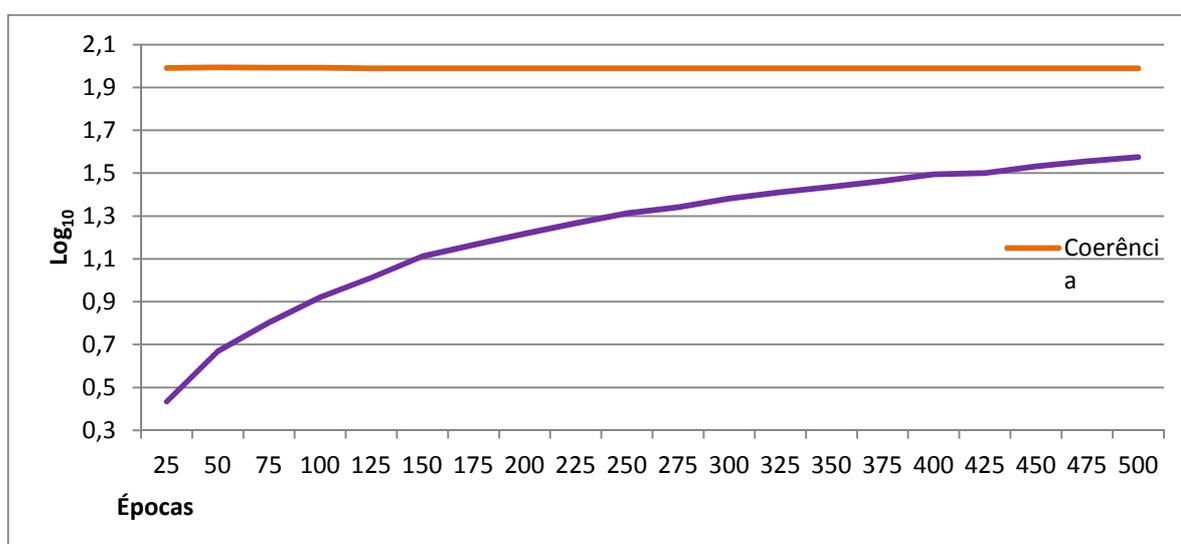
Gráfico 3 Treinamento da Base 2 – Níveis ocultos: coerência x tempo



Fonte: Autoria própria, 2011.

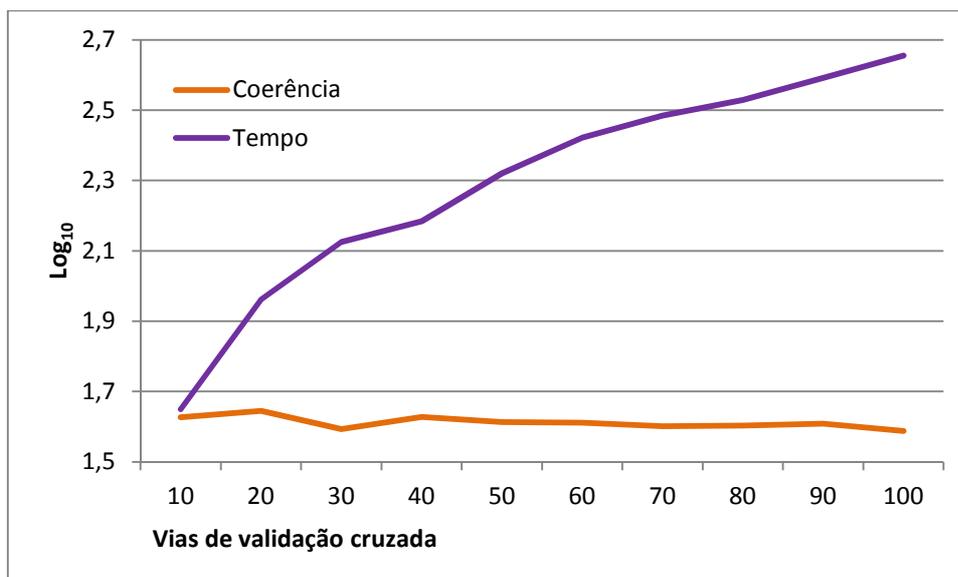
Já o Gráfico 3 (acima) apresenta os resultados dos experimentos executados na Base 2 utilizando rede MLP variando a quantidade de níveis ocultos de 1 a 100, demonstrando os resultados em base logarítmica da relação coerência *versus* tempo de treinamento. Abaixo, o Gráfico 4 demonstra os resultados da mesma relação da Base 2 no treinamento da MLP, também em base logarítmica, variando a quantidade de épocas de retro propagação de 25 a 500 em intervalos de 25 épocas.

Gráfico 4 Treinamento da Base 2 – Épocas: coerência x tempo



Fonte: Autoria própria, 2011.

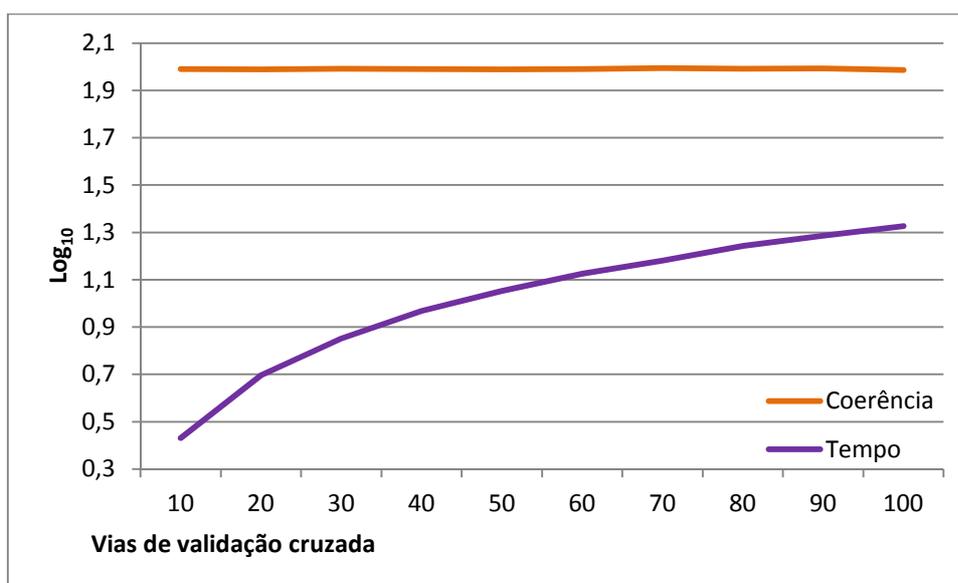
Gráfico 5 Treinamento da Base 1 – Vias: coerência x tempo



Fonte: Autoria própria, 2011.

Os Gráficos 5 e 6 apresentam os resultados dos experimentos de treinamento das MLPs das Bases 1 e 2, respectivamente, variando a quantidade de vias de validação cruzada de 10 a 100 em intervalos de 10 vias, com a relação entre coerência e tempo de treinamento também em base logarítmica.

Gráfico 6 Treinamento da Base 2 – Vias: coerência x tempo



Fonte: Autoria própria, 2011.

Pôde-se chegar a um valor coerente aproximado para o algoritmo de redes neurais MLP, baseando os parâmetros nos resultados dos experimentos anteriores, de 44,1824% para a Base 1 e 98,6842 % para a Base 2. Comparando os resultados com o algoritmo de árvores de decisão J48, a rede neural foi muito menos eficaz para Base 1 e mais eficaz para Base 2. A Tabela 2 e o Gráfico 7 apresentam a comparação entre os resultados obtidos nas duas bases utilizando ambos os algoritmos.

A árvore de decisão apresenta melhor desempenho, sua execução foi constantemente inferior a 30 milissegundos para ambas as bases, já a rede neural variou seu treinamento entre 44,6 segundos e 7 minutos e 11 segundos para a Base 1 e entre 2,6 segundos e 21,2 segundos para a Base 2, confirmando o algoritmo J48 como mais adequado para a base de dados da primeira empresa e a rede neural MLP mais eficaz e eficiente para tratar as características de classificação e categorização da segunda empresa.

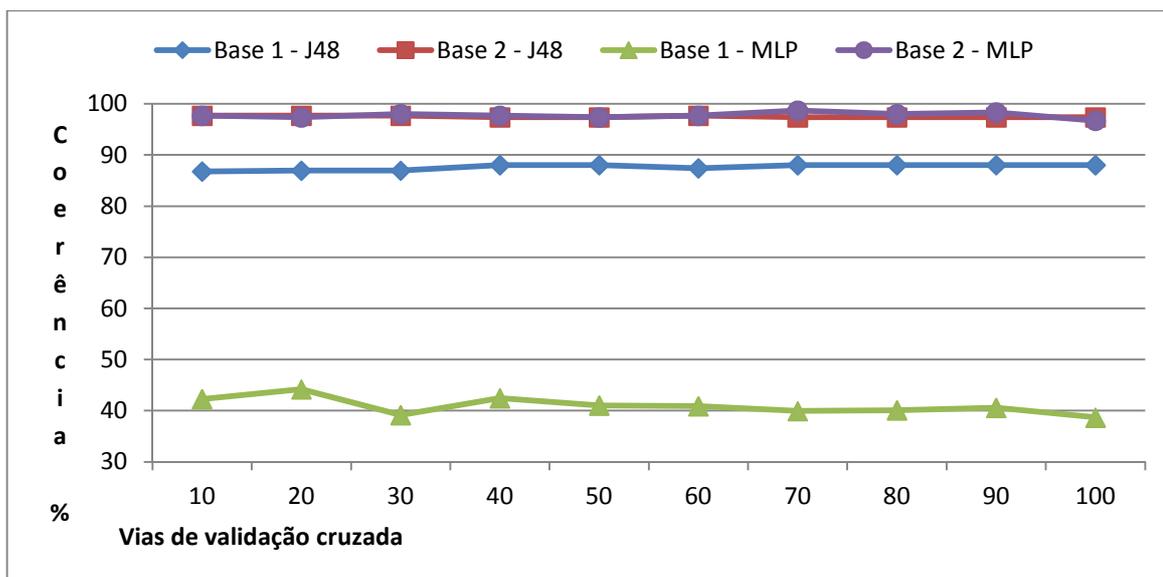
Tabela 2 Comparativo de resultados de coerência dos experimentos

Validação Cruzada	Base 1 - J48	Base 2 - J48	Base 1 - MLP	Base 2 - MLP
10	86,7925	97,6974	42,2956	97,6974
20	86,9497	97,6974	44,1824	97,3684
30	86,9497	97,6974	39,1509	98,0263
40	88,0503	97,3684	42,4528	97,6974
50	88,0503	97,3684	41,0377	97,3684
60	87,4214	97,6974	40,8805	97,6974
70	88,0503	97,3684	39,9371	98,6842
80	88,0503	97,3684	40,0943	98,0263
90	88,0503	97,3684	40,566	98,3553
100	88,0503	97,3684	38,6792	96,7105

Fonte: Autoria própria, 2011.

A Tabela 2 e o Gráfico 7 apresentam a comparação entre os resultados obtidos nas duas bases utilizando ambos os algoritmos. Confirmando a ineficácia do treinamento da rede MLP para a Base 1.

Gráfico 7 Comparativo de resultados de coerência dos experimentos



Fonte: Autoria própria, 2011.

4.6 Considerações finais

Neste penúltimo capítulo foram apresentadas as características das informações utilizadas, como os algoritmos para avaliação foram selecionados, a metodologia e transformações nos dados, o ambiente e parâmetros de execução dos experimentos e seus resultados analisados. No próximo capítulo será concluído o raciocínio iniciado na introdução e desenvolvido durante todo o trabalho.

5 CONCLUSÃO

Tendo em vista as experimentações e os resultados obtidos, tem-se a indicação positiva sobre o objetivo principal do trabalho, podendo auxiliar na tomada de decisões com relação à definição e/ou alteração de uma classificação e categorização dentro do processo de Gerenciamento de Incidentes modelado a partir da ITIL para ambas as empresas. Aparentemente o auxílio é mais significativo em uma base mais voltada à tecnologia, onde se pode ter dúvidas e uma maior dificuldade na identificação e caracterização do problema do usuário.

Identificou-se que as técnicas não se comportam de forma semelhante para a classificação e categorização das duas empresas, a base da primeira empresa possui características que desfavorecem a utilização de redes neurais *multi-layer perceptron*, enquanto que as propriedades da base desenhada pela segunda empresa são aderentes em ambos os algoritmos, com acanhado rendimento extra na rede neural MLP.

Seguem em apêndices os arquivos ARFF utilizados na inferência das duas bases para Weka, podendo estes serem utilizados em novos experimentos e trabalhos futuros. O conhecimento agregado para redução no tempo de treinamento da rede neural MLP e consumo computacional através da codificação dos dados em números primos mostrou-se incrivelmente eficiente (dado o percentual do ganho obtido).

O retorno das empresas sobre os resultados obtidos neste trabalho foram muito positivos. A primeira empresa criou vínculo entre classificação e o primeiro nível de categorização e está revisando sua classificação e categorização com o intuito de reduzir a possibilidade de desvios. Além desta revisão, a primeira empresa está avaliando a integração da classe de árvores de decisão J48 do Weka em seu software de gerenciamento de incidentes. A segunda empresa já acreditava que sua classificação e categorização eram eficientes, mas ainda não haviam estruturado um indicador, sendo este trabalho muito bem vindo, apresentando um valor numérico confirmando a sensação da equipe e gestão.

Como outro trabalho futuro, sugere-se avaliar outros algoritmos e técnicas de Inteligência Artificial e métodos de inferência para este problema, como Algoritmos

Genéticos, e Máquinas de Núcleo que podem também fornecer soluções interessantes para o problema.

REFERÊNCIAS

- ALMEIDA, L. M. et al. Uma ferramenta para extração de padrões, Palmas, 2004. Disponível em: <<http://143.54.31.10/reic/edicoes/2003e4/cientificos/UmaFerramentaParaExtracaoDePadroes.pdf>>. Acesso em: Setembro 2011.
- BRÜGGEMANN, E. S.; PORTO, R. M. Uma proposta ontológica para um sistema de gestão de versionamento de modelo ITIL, Brasília, Março 2006. Disponível em: <<http://monografias.cic.unb.br/dspace/bitstream/123456789/37/1/MonografiaEduardoRicardo.pdf>>. Acesso em: Setembro 2011. Monografia apresentada como requisito parcial para conclusão do Bacharelado em Ciência da Computação.
- CALDWELL, C. Lists of small primes. **The Prime Pages**, 2009. Disponível em: <<http://primes.utm.edu/>>. Acesso em: Agosto 2011.
- CALVI, C. Z. Gerenciamento de serviços de TI e modelagem do processo de configuração ITIL em uma plataforma de serviços sensíveis a contexto, Vitória, Novembro 2007. Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Informática da Universidade Federal do Espírito Santo como requisito parcial para obtenção do grau de Mestre em Informática.
- CARVALHO, A. P. D. L. F. D. Redes Neurais Artificiais. **Department of Computer Science University of São Paulo**, 2003. Disponível em: <<http://www.icmc.usp.br/~andre/>>. Acesso em: Maio 2011.
- GAMA, J. Árvores de Decisão, 2002. Disponível em: <http://www.liacc.up.pt/~jgama/Aulas_ECD/arv.pdf>. Acesso em: Setembro 2011.
- GUARDA, Á. Aprendizado de Máquina: Árvore de Decisão Indutiva. Disponível em: <<http://www.decom.ufop.br/prof/guarda/CIC250/ArvoreDecisaoIndutiva.pdf>>. Acesso em: Setembro 2011.
- ILLUMNA FOUNDATION CERTIFICATE. **ITIL Foundations**. 1.3. ed. São Paulo: Exin, 2006.
- ILLUMNA FOUNDATION CERTIFICATE. **ITILv3 Foundations Bridging in IT Service Management**. 1.2. ed. São Paulo: Exin, 2007.
- MENEZES, H. N. Avaliação do nível de maturidade da governança de tecnologia da informação: Estudo de caso em indústrias de grande porte, Fortaleza, 2005. Dissertação submetida ao corpo docente do Curso de Mestrado em Informática Aplicada da Universidade de Fortaleza como requisito parcial para a obtenção do título de Mestre em Ciência da Computação.
- MILIES, F. C. P.; COELHO, S. P. **Números: Uma Introdução à Matemática**. 3ª. ed. São Paulo: Editora da Universidade de São paulo, 2003.
- MURALIKRISHNA, A. et al. Busca de Regras para Previsão de Tempestades Geomagnéticas Utilizando Técnica de Mineração de Dados, 2005. Disponível em:

<<http://hermes2.dpi.inpe.br:1905/col/dpi.inpe.br/hermes2%401905/2005/10.03.15.35/doc/VWorcap.pdf>>. Acesso em: Maio 2011.

NIEVOLA, J. C. Redes Neurais Artificiais. **Pontifícia Universidade Católica do Paraná**, 1999. Disponível em: <<http://www.ppgia.pucpr.br/~nievola/>>. Acesso em: Maio 2011.

ONODA, M.; EBECKEN, N. F. F. Implementação em Java de um Algoritmo de Árvore de Decisão Acoplado a um SGBD Relacional, 2001. Disponível em: <<http://www.lbd.dcc.ufmg.br:8080/colecoes/sbbd/2001/004.pdf>>. Acesso em: Setembro 2011.

POZO, A. T. R. Algoritmo de Aprendizado de Máquina. Disponível em: <<http://www.inf.ufpr.br/aurora/tutoriais/arvoresdecisao/>>. Acesso em: Setembro 2011.

RHINO CONSULTING. **Rhino Get Started! Transição e Operação**. Porto Alegre: [s.n.], 2010.

RUSSELL, S.; NORVIG, P. **Inteligência Artificial**. [S.l.]: Campus, 2004.

SANTOS, R. A. D. Avaliação do risco associado a mudanças de TI com aplicação para a atividade de priorização, Campina Grande, Agosto 2007.

SANTOS, R. D. C. D. Data mining: Conceitos e Algoritmos, 18 Agosto 2005a. Disponível em: <<http://www.lac.inpe.br/~rafael.santos/Docs/CursoInverno/2005/ci.dm.pdf>>.

SANTOS, R. D. C. D. Weka na Munheca - Um guia para uso do Weka em scripts e integração com aplicações em Java, 2005b. Disponível em: <<http://www.lac.inpe.br/~rafael.santos/Docs/CAP359/2005/weka.pdf>>. Acesso em: Janeiro 2011.

SIMÕES, A. D. S.; COSTA, A. H. R. **Redes neurais pulsadas de base radial com treinamento não supervisionado**: Um estudo da capacidade de agrupamento para a visão robótica. VII SBAI / II IEEE LARS. São Luis: [s.n.]. 2005.

TATIBANA, K.; CASSIA, D. Homepage de Redes Neurais. **Departamento de Informática da Universidade Estadual de Maringá**, 2000. Disponível em: <<http://www.din.uem.br/>>. Acesso em: Junho 2011.

THOMÉ, A. C. G. Inteligência Computacional Aula 11 – MLP (MultiLayer Perceptron). **Universidade Federal do Rio de Janeiro Núcleo de Computação Eletrônica**, 2003. Disponível em: <<http://equipe.nce.ufrj.br/thome>>. Acesso em: Setembro 2011.

WITTEN, I. H.; FRANK, E. **Data mining**: Practical machine learning tools and techniques. 2^a. ed. San Francisco: [s.n.], 2005.

APÊNDICE A – TEOREMA FUNDAMENTAL DA ARITMÉTICA

O Teorema Fundamental da Aritmética destaca a importância dos números primos na Teoria dos Números, eles desempenham um papel análogo ao dos átomos na constituição da matéria, todos os outros números podem ser obtidos através de produtos de números primos para todo inteiro diferente de -1, 1 e 0.

Um inteiro p diz-se primo se tem exatamente dois divisores inteiros positivos, 1 e $|p|$ e, um número diferente de -1, 1 e 0 e que não é primo diz-se composto.

Teorema (MILIES e COELHO, 2003):

Seja a um inteiro diferente de 1,0 e -1. Então, existem primos positivos $p_1 < p_2 < \dots < p_y$ e inteiros positivos n_1, n_2, \dots, n_y tais que $a = Ep_1^{n_1} \dots Ep_y^{n_y}$ em que $E = \pm 1$, conforme a seja positivo ou negativo. Além disso, esta decomposição é única.

Em função destas características, os números primos são base da codificação de dados, pois permitem uma decodificação única e direta.

APENDICE B – BASE 1

@relation Base1

@attribute classificacao {2,3,5}

@attribute categoria1 {2,3,5,7,11,13,17,19,23,29,31}

@attribute categoria2

{2,3,5,7,11,13,17,19,23,29,31,37,41,43,47,53,59,61,67,71,73,79,83,89,97,101,103,107,109,113,127,131,137,139,149,151,157,163,167,173,179,181,191,193,197,199,211,223,227,229,233,239,241,251,257,263,269,271,277,281,283,293,307,311,313,317,331,337,347,349,353,359,367,373,379,383,389,397,401,409,419,421,431}

@attribute categoria3

{2,3,5,7,11,13,17,19,23,29,31,37,41,43,47,53,59,61,67,71,73,79,83,89,97,101,103,107,109,113,127,131,137,139,149,151,157,163,167,173,179,181,191,193,197,199,211,223,227,229,233,239,241,251,257,263,269,271,277,281,283,293,307,311,313,317,331,337,347,349,353,359,367,373,379,383,389,397,401,409,419,421,431,433,439,443,449,457,461,463,467,479,487,491,499,503,509,521,523,541,547,557,563,569,571,577,587,593,599,601,607,613,617,619,631,641,643,647,653,659,661,673,677,683,691,701,709,719,727,733,739,743,751,757,761,769,773,787,797,809,811,821,823,827,829,839,853,857,859,863,877,881,883,887,907,911,919,929,937,941,947,953,967,971,977,983,991,997,1009,1013,1019,1021,1031,1033,1039,1049,1051,1061,1063,1069,1087,1091,1093,1097,1103,1109,1117,1123,1129,1151,1153,1163,1171,1181,1187,1193,1201,1213,1217,1223,1229,1231,1237,1249,1259,1277,1279,1283,1289,1291,1297,1301,1303,1307,1319,1321,1327,1361,1367,1373,1381,1399,1409,1423,1427,1429,1433,1439,1447,1451,1453,1459,1471,1481,1483,1487,1489,1493,1499,1511,1523,1531,1543,1549,1553,1559,1567,1571,1579,1583,1597,1601,1607,1609,1613,1619,1621,1627,1637,1657,1663,1667,1669,1693,1697,1699,1709,1721,1723,1733,1741,1747,1753,1759,1777,1783,1787,1789,1801,1811,1823,1831,1847,1861,1867,1871,1873,1877,1879,1889,1901,1907,1913,1931,1933,1949,1951,1973,1979,1987,1993,1997,1999,2003,2011,2017,2027,2029,2039,2053,2063,2069,2081,2083,2087,2089,2099,2111,2113,2129,2131,2137,2141,2143,2153,2161,2179,2203,2207,2213,2221,2237,2239,2243,2251,2267,2269,2273,2281,2287,2293,2297}

@data

5,5,73,643
 5,5,73,929
 5,5,73,1301
 5,5,103,137
 5,5,103,79
 5,5,103,2053
 3,5,127,1063
 3,5,127,1049
 3,5,127,1471
 5,5,167,1511
 3,13,173,1559
 3,13,173,1567
 3,13,173,2087
 5,13,173,379
 5,13,173,401

5,13,173,757
5,13,173,1013
5,13,173,1019
5,13,173,1511
5,13,173,1583
5,13,173,1637
5,13,173,1873
3,13,173,2069
5,13,173,2113
3,13,179,1481
3,13,179,1483
3,13,179,1499
3,13,179,1543
3,13,179,1559
3,13,179,1567
3,13,179,2087
5,13,179,401
5,13,179,1013
5,13,179,1019
5,13,179,1511
2,13,179,1511
5,13,179,1583
5,13,179,1873
5,13,179,2113
3,13,223,617
5,13,223,379
5,13,73,643
5,13,223,809
5,13,223,829
5,13,223,1019
3,13,223,1063
5,13,223,1511
5,13,223,1657
5,13,223,1667
5,13,223,1873
3,29,17,97
3,29,43,227
3,29,43,367
3,29,43,617
3,29,181,1129
3,29,181,233
3,29,181,311
3,29,17,1453
3,29,17,1447
3,29,17,1451
5,13,223,1877
3,13,173,1553
3,13,179,1553
3,29,199,887
5,29,73,7

5,29,199,7
5,13,179,2113
5,5,73,7
5,13,73,7
5,29,73,643
5,13,73,929
5,29,73,929
5,13,73,1301
5,29,199,1637
5,29,199,271
3,13,223,1831
3,13,223,1847
3,13,223,1823
3,13,173,233
3,13,173,1549
3,13,179,233
3,13,179,1549
3,13,223,367
3,13,223,1543
3,29,67,433
3,29,191,1499
3,29,191,1493
3,29,191,1487
3,29,11,83
3,29,191,1129
3,29,191,311
5,13,179,389
3,29,191,1489
5,13,173,2081
5,13,179,2081
5,13,223,2081
3,13,173,1811
5,29,191,1511
2,29,191,1511
5,29,199,907
5,29,199,2113
5,13,179,397
5,13,173,397
3,13,223,1319
3,13,173,1319
3,29,211,1049
3,2,157,1913
5,2,233,89
3,2,233,89
5,2,233,631
3,2,233,631
2,2,233,631
5,2,233,683
3,2,233,683
3,2,233,877

5,2,233,877
2,2,233,877
5,2,233,1039
3,2,233,1039
5,2,233,1087
3,2,233,1087
5,2,233,1297
3,2,233,1297
3,2,281,1889
3,2,293,877
3,3,3,2
5,3,3,2
2,3,3,2
3,3,3,3
5,3,3,3
2,3,3,3
3,3,3,421
5,3,3,421
2,3,3,421
3,3,3,2129
5,3,3,2129
2,3,3,2129
3,3,3,2297
5,3,3,2297
2,3,3,2297
3,3,47,5
5,3,47,5
2,3,47,5
3,3,47,1021
3,3,47,1031
3,3,47,1033
2,3,47,1033
3,3,47,1187
3,3,47,1193
3,3,47,1217
3,3,83,1103
2,3,83,1103
3,3,83,1109
5,3,83,1109
3,3,83,1531
5,3,83,1531
2,3,83,1531
3,3,83,2089
5,3,83,2089
3,3,83,2161
3,3,83,2179
3,3,83,1097
3,3,97,691
2,3,97,691
3,3,97,701

3,3,97,719
3,3,97,727
3,3,97,733
3,3,97,1237
3,3,97,1259
2,3,97,1259
3,3,97,1321
3,3,107,919
3,3,107,1151
3,3,107,1153
3,3,109,503
3,3,109,919
3,3,109,1163
3,3,109,1231
3,3,113,19
2,3,113,19
3,3,113,31
2,3,113,31
3,3,113,113
5,3,113,113
2,3,113,113
3,3,113,431
5,3,113,431
2,3,113,431
3,3,113,457
3,3,113,859
5,3,113,859
2,3,113,859
3,3,113,919
3,3,113,1151
3,3,113,1201
3,3,113,1213
3,3,113,1223
3,3,113,1367
3,3,113,2293
5,3,113,2293
2,3,113,2293
3,3,131,593
3,3,131,1187
3,3,131,1151
3,3,131,1237
3,3,131,1931
3,3,139,521
3,3,139,577
3,3,139,601
3,3,139,607
3,3,139,919
3,3,139,1231
3,3,139,2089
3,3,151,509

3,3,151,587
3,3,151,599
3,3,151,919
3,3,163,523
3,3,163,541
3,3,163,547
3,3,163,557
3,3,163,563
3,3,163,569
3,3,251,1459
3,3,251,571
3,3,251,641
3,3,251,919
3,3,251,1171
3,3,251,1217
3,3,251,1621
3,3,251,2099
3,3,269,11
5,3,269,11
2,3,269,11
3,3,269,17
5,3,269,17
2,3,269,17
3,3,269,173
5,3,269,173
2,3,269,173
3,3,269,229
5,3,269,229
2,3,269,229
3,3,269,439
5,3,269,439
2,3,269,439
3,3,269,883
2,3,269,883
3,3,269,911
5,3,269,911
2,3,269,911
3,3,269,1069
5,3,269,1069
2,3,269,1069
3,3,269,1291
5,3,269,1291
2,3,269,1291
3,3,269,1303
5,3,269,1303
2,3,269,1303
3,3,269,1523
5,3,269,1523
2,3,269,1523
3,3,269,1901

5,3,269,1901
2,3,269,1901
3,3,269,2243
5,3,269,2243
2,3,269,2243
3,31,19,233
3,31,19,449
3,31,19,1129
3,31,19,1229
3,31,19,1249
3,31,19,1399
3,31,19,2273
3,2,83,709
5,2,83,709
2,2,83,709
3,2,83,1093
5,2,83,1093
2,2,83,1093
3,2,83,1103
5,2,83,1103
2,2,83,1103
3,2,83,257
5,2,83,257
2,2,83,257
3,2,83,1277
5,2,83,1277
2,2,83,1277
3,2,233,1427
5,2,233,1427
3,2,233,2203
5,2,233,2203
3,3,113,739
2,3,113,739
5,2,29,1871
2,2,29,1871
3,2,233,359
5,2,233,359
3,3,113,1579
5,3,113,1579
2,3,113,1579
3,3,13,743
3,3,83,1423
3,3,83,1409
5,3,83,1409
3,2,229,13
5,2,229,13
5,2,5,1117
5,2,5,37
5,2,5,41
3,2,233,2063

5,2,233,2063
3,3,83,1303
5,3,83,1303
2,3,83,1303
3,3,269,1381
5,3,269,1381
2,3,269,1381
3,2,101,109
5,2,101,109
3,3,113,1619
2,3,113,1619
5,3,113,1619
3,2,233,181
5,2,233,181
3,3,113,463
3,2,283,461
3,2,283,373
3,2,283,751
3,2,157,1907
3,2,233,2267
5,2,233,2267
3,3,113,197
2,3,113,197
3,2,233,1303
5,2,233,1303
2,2,233,1303
3,2,281,659
5,2,227,419
3,2,227,283
3,2,227,127
3,3,227,1373
3,3,227,5
3,3,227,2213
5,2,233,1979
2,2,233,1979
3,2,233,1979
3,3,3,2281
2,3,3,2281
3,2,233,211
5,2,233,211
5,2,5,43
3,29,181,1571
3,29,191,1571
5,29,73,1301
5,7,311,809
5,7,311,1637
5,7,317,131
5,7,349,757
5,7,349,1637
5,7,373,271

5,7,373,887
5,7,373,907
5,7,373,1637
5,7,379,757
5,7,389,1637
5,7,89,1993
5,7,89,53
5,7,89,2011
5,7,89,2027
5,7,89,2017
5,7,89,2003
5,7,89,1999
5,7,89,2029
5,7,89,2039
5,7,197,761
5,7,197,1787
5,7,409,757
3,29,307,1429
3,29,307,613
3,29,307,619
3,29,421,307
3,29,431,307
3,3,227,281
5,17,23,673
5,17,23,677
5,17,23,1871
5,17,37,37
5,17,47,41
5,17,47,59
5,17,47,61
5,17,47,71
5,17,47,349
5,17,47,191
5,17,47,383
5,17,47,29
5,17,47,947
5,17,47,953
5,17,47,967
5,17,47,971
5,17,47,983
5,17,47,1009
5,17,47,1697
5,17,47,1699
5,17,47,1709
5,17,47,1669
5,17,47,1747
5,17,47,1777
5,17,47,1693
5,17,47,1789
5,17,47,2137

5,17,59,67
5,17,59,941
5,17,59,1741
5,17,59,383
5,17,59,1283
5,17,59,1289
5,17,53,971
5,17,53,977
5,17,53,1009
5,17,53,991
5,17,53,997
5,17,53,1723
5,17,53,1721
5,17,53,1733
5,17,53,1777
5,17,53,1753
5,17,53,1759
2,17,61,877
2,17,61,1327
2,17,61,1951
2,17,61,1291
2,17,61,1307
2,17,61,2111
2,17,61,1879
2,17,61,1279
2,17,61,1933
5,17,89,277
5,17,89,881
5,17,113,29
5,17,113,23
5,17,113,167
5,17,113,193
5,17,113,263
5,17,113,647
5,17,113,277
5,17,113,331
5,17,113,383
5,17,113,1801
5,17,113,317
5,17,113,313
5,17,113,277
5,17,113,1619
2,17,113,1619
5,17,193,661
5,17,257,107
5,17,257,149
5,17,257,653
5,17,409,1997
5,17,409,1511
5,17,409,1511

5,17,311,2113
5,17,313,757
5,17,313,1511
5,17,313,1583
5,17,317,1061
5,17,317,1511
5,17,317,1877
5,17,317,2087
5,17,317,2221
5,17,331,1511
3,17,331,2087
5,17,337,757
5,17,337,1877
5,17,347,757
5,17,347,787
5,17,347,823
5,17,347,1637
3,17,347,2087
5,17,347,2113
5,17,349,2113
5,17,353,7
5,17,353,643
5,17,353,929
5,17,353,1301
5,17,359,1637
5,17,359,379
5,17,359,787
5,17,359,811
5,17,359,827
5,17,359,1511
5,17,359,2083
5,17,359,2113
5,17,367,379
5,17,367,643
5,17,367,827
5,17,367,1511
5,17,367,1867
5,17,367,2083
5,17,373,7
5,17,373,1949
5,17,373,2113
5,17,379,1877
5,17,383,199
5,17,383,1627
5,17,383,1877
5,17,383,2237
5,17,389,1877
5,17,397,1877
5,17,401,1877
5,17,401,7

3,29,199,1607
3,29,67,1433
3,29,67,1439
3,29,79,2239
5,17,257,2153
5,17,257,2143
5,17,257,2141
5,17,47,1783
5,17,257,139
5,17,113,1861
5,17,113,2131
5,17,113,277
5,29,43,2143
5,7,197,857
5,17,23,293
5,17,359,1123
5,17,89,1051
5,17,257,331
5,17,89,1637
5,7,197,1663
3,3,83,1361
5,3,83,1361
2,3,83,1361
5,17,257,347
3,3,251,467
3,13,173,863
3,13,179,863
3,13,223,863
3,29,199,251
5,29,199,251
3,29,199,1597
5,7,227,353
5,7,227,1609
5,17,89,821
5,13,7,2251
3,2,101,1601
5,2,101,1601
2,2,101,1601
5,17,311,809
5,17,311,1049
5,7,227,1613
5,7,239,163
5,7,239,157
5,7,239,151
5,7,197,797
5,17,89,1987
3,2,277,937
3,2,83,2207
5,2,83,2207
5,7,241,337

5,7,89,773
5,7,263,853
5,17,263,149
5,7,241,73
3,2,229,409
5,2,229,409
2,2,229,409
3,2,229,443
5,2,229,443
2,2,229,443
3,3,269,1091
3,2,419,487
3,2,419,479
3,2,419,491
3,2,233,1069
5,2,233,1069
2,2,233,1069
5,2,71,223
3,3,269,1979
5,3,269,1979
2,3,269,1979
2,19,271,661
3,19,271,661
5,7,241,103
5,17,47,277
5,17,251,277
2,17,251,277
3,3,113,2273
3,3,13,47
2,17,61,2287
3,2,281,257
3,2,281,181
3,2,227,499
5,7,241,769
3,2,233,2269
2,2,233,2269
3,3,211,1181
3,3,211,269
5,3,211,277
2,3,211,277
3,3,113,179
5,3,113,179
3,2,233,1879
5,2,233,1879
3,3,269,239
2,3,269,239
5,3,269,239
3,31,41,743
3,23,2,661
5,23,2,661

3,3,47,1931
5,17,89,839
5,17,31,277
3,3,257,101
5,3,257,101
2,3,257,101
3,3,257,241
5,3,257,241
2,3,257,241
3,3,257,1879
5,3,257,1879
2,3,257,1879
3,3,257,1973
5,3,257,1973
2,3,257,1973
3,23,149,661
5,23,149,661
2,23,149,661
5,11,137,1093
5,11,137,1277
5,11,137,709

APÊNDICE C – BASE 2

@relation Base1

@attribute classificacao {2,3,5,7}

@attribute categoria1 {2,3,5,7,11,13,17,19,23}

@attribute categoria2 {2,3,5,7,11,13,17,19,23,29,31,37,41,43,47,53,59,61,67,71}

@attribute categoria3

{2,3,5,7,11,13,17,19,23,29,31,37,41,43,47,53,59,61,67,71,73,79,83,89,97,101,103,107,109,113,127,131,137,139,149,151,157,163,167,173,179,181,191,193,197,199,211,223,227,229,233,239,241,251,257,263,269,271,277,281,283,293,307,311,313,317,331,337,347,349,353,359,367,373,379,383,389,397,401,409,419,421,431,433,439,443,449,457,461,463,467,479,487,491,499,503,509,521,523,541,547,557,563,569,571,577,587,593,599,601,607,613,617,619,631,641,643,647,653,659,661,673,677,683,691,701,709,719,727,733,739,743,751}

@data

5,19,61,311
5,19,61,331
5,19,61,337
5,19,61,409
5,19,61,463
5,19,61,631
5,19,61,251
5,19,61,647
5,19,61,257
5,19,61,23
5,19,67,691
5,19,67,743
5,19,61,41
5,19,61,73
5,19,61,113
5,19,61,127
5,19,67,443
5,19,67,457
5,19,67,499
5,19,67,569
5,19,67,577
5,19,67,601
5,19,59,317
5,19,59,347
5,19,59,419
5,19,59,491
5,19,67,367
5,19,67,439
5,19,59,19
5,19,59,59
5,19,59,131
5,19,59,151

5,19,59,181
5,19,59,269
5,19,47,359
5,19,47,397
5,19,47,401
5,19,47,509
5,19,53,607
5,19,53,613
5,19,47,31
5,19,47,79
5,19,47,89
5,19,47,223
5,19,47,263
5,19,47,283
5,19,67,241
5,19,67,271
5,19,67,383
5,19,67,431
5,19,47,461
5,19,47,547
5,19,67,137
5,19,67,149
5,19,67,167
5,19,67,193
5,19,67,199
5,19,67,211
5,19,59,521
5,19,59,571
5,19,59,659
5,19,59,661
5,19,59,719
5,19,67,97
5,19,43,557
5,19,43,587
5,19,43,599
5,19,43,641
5,19,43,727
5,19,47,11
5,19,43,179
5,19,43,227
5,19,43,239
5,19,43,449
5,19,43,523
5,19,43,541
5,19,41,277
5,19,41,433
5,19,41,593
5,19,41,619
5,19,43,157
5,19,43,163

5,19,37,701
5,19,37,739
5,19,41,61
5,19,41,67
5,19,41,107
5,19,41,233
5,19,37,13
5,19,37,37
5,19,37,53
5,19,37,503
5,19,37,617
5,19,37,683
5,19,23,173
5,19,23,751
5,19,29,43
5,19,29,83
5,19,29,197
5,19,29,373
5,19,37,277
5,19,37,349
5,19,37,353
5,19,37,379
5,19,37,487
5,19,23,47
5,19,61,643
5,19,37,71
5,19,37,101
5,19,37,109
5,19,37,191
5,19,37,229
2,3,11,563
3,3,11,563
2,3,11,467
3,3,11,467
2,3,11,389
3,3,11,389
2,3,11,599
3,3,11,599
2,11,37,109
2,11,37,229
2,11,37,53
2,11,37,353
2,11,37,71
2,11,37,13
2,11,41,61
2,11,41,107
2,11,41,67
2,11,41,233
2,11,41,433
2,11,41,277

2,11,41,593
2,11,41,619
2,11,37,379
2,11,37,701
2,11,37,101
2,11,37,683
2,11,37,349
2,11,37,503
2,11,37,37
2,11,37,617
2,11,37,739
2,11,37,277
2,11,37,191
2,11,37,487
2,13,71,29
2,13,71,281
2,23,31,241
2,23,31,271
2,23,31,431
2,23,31,439
2,23,31,443
2,23,31,457
2,23,31,167
2,23,31,193
2,23,31,199
2,23,31,211
2,23,2,173
2,23,2,751
2,23,2,47
2,23,31,97
2,23,31,137
2,23,31,149
2,23,31,367
2,23,2,653
2,23,3,43
2,23,3,83
2,23,3,197
2,23,3,373
2,23,31,499
2,23,31,577
2,23,31,601
2,23,31,691
2,23,31,743
2,23,31,383
5,17,41,107
5,17,41,233
5,17,41,277
5,17,41,433
5,17,41,593
5,17,41,619

5,17,61,647
5,17,61,257
5,17,61,23
5,17,59,131
5,17,59,151
5,17,59,181
5,17,59,269
5,17,59,317
5,17,67,367
5,17,37,701
5,17,37,739
5,17,41,61
5,17,41,67
5,17,59,19
5,17,59,59
5,17,37,353
5,17,37,379
5,17,37,487
5,17,37,503
5,17,37,617
5,17,37,683
5,17,59,521
5,17,59,571
5,17,59,659
5,17,59,661
5,17,59,719
5,17,67,97
5,17,61,631
5,17,61,643
5,17,47,547
5,17,59,347
5,17,59,419
5,17,59,491
5,17,61,251
5,17,61,311
5,17,61,331
5,17,61,337
5,17,61,409
5,17,61,463
5,17,67,691
5,17,67,743
5,17,61,41
5,17,61,73
5,17,61,113
5,17,61,127
5,17,67,443
5,17,67,457
5,17,67,499
5,17,67,569
5,17,67,577

5,17,67,601
5,17,67,211
5,17,67,241
5,17,67,271
5,17,67,383
5,17,67,431
5,17,67,439
5,17,37,349
5,17,67,137
5,17,67,149
5,17,67,167
5,17,67,193
5,17,67,199
5,17,37,71
5,17,37,101
5,17,37,109
5,17,37,191
5,17,37,229
5,17,37,277
5,17,29,83
5,17,29,197
5,17,29,373
5,17,37,13
5,17,37,37
5,17,37,53
5,17,53,607
5,17,53,613
5,17,23,47
5,17,23,173
5,17,23,751
5,17,29,43
5,17,47,263
5,17,47,283
5,17,47,359
5,17,47,397
5,17,47,401
5,17,47,509
5,17,43,727
5,17,47,11
5,17,47,31
5,17,47,79
5,17,47,89
5,17,47,223
5,17,43,523
5,17,43,541
5,17,43,557
5,17,43,587
5,17,43,599
5,17,43,641
5,17,43,157

5,17,43,163
5,17,43,179
5,17,43,227
5,17,43,239
5,17,43,449
5,7,13,479
5,7,19,7
5,7,19,17
5,7,19,313
5,7,19,479
5,7,19,709
5,7,19,139
7,2,17,421
7,2,17,293
7,2,17,103
7,2,17,733
7,2,7,307
7,2,7,677
7,2,7,673
7,5,5,3
7,5,5,2
7,5,5,5