



**Pós-Graduação em Ciência da Computação**

**“UMA INVESTIGAÇÃO SOBRE O PROCESSO MIGRATÓRIO  
PARA A PLATAFORMA DE COMPUTAÇÃO EM NUVEM NO  
BRASIL”**

**Por**

***Hilson Barbosa da Silva***

**Dissertação de Mestrado Profissional**



Universidade Federal de Pernambuco  
posgraduacao@cin.ufpe.br  
[www.cin.ufpe.br/~posgraduacao](http://www.cin.ufpe.br/~posgraduacao)

RECIFE, 2016



Universidade Federal de Pernambuco

CENTRO DE INFORMÁTICA

PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

***Hilson Barbosa da Silva***

***“UMA INVESTIGAÇÃO SOBRE O PROCESSO  
MIGRATÓRIO PARA A PLATAFORMA DE  
COMPUTAÇÃO EM NUVEM NO BRASIL”***

Este trabalho foi apresentado à Pós-Graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Mestre

***ORIENTADOR(A): Prof. Vinicius Cardoso Garcia***

***COORIENTADOR(A): Prof. Thiago Carvalho de Sousa***

RECIFE, 2016

Catálogo na fonte  
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

S586i Silva, Hilson Barbosa da  
Uma investigação sobre o processo migratório para a plataforma de  
computação em nuvem no Brasil / Hilson Barbosa da Silva. – 2016.  
109 f.: il., fig., tab.  
  
Orientador: Vinicius Cardoso Garcia.  
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CIn,  
Ciência da Computação, Recife, 2016.  
  
Inclui referências e apêndices.  
  
1. Ciência da computação. 2. Inteligência artificial. 3. Computação em  
nuvem. I. Garcia, Vinicius Cardoso (orientador). II. Título.  
  
004 CDD (23. ed.) UFPE- MEI 2016-068

HILSON BARBOSA DA SILVA

**UMA INVESTIGAÇÃO DO PROCESSO MIGRATÓRIO PARA  
A PLATAFORMA DE COMPUTAÇÃO EM NUVEM NO BRASIL**

Dissertação apresentado à pós-graduação em  
Ciência da Computação do Centro de Informática da  
Universidade Federal de Pernambuco – UFPE, como  
requisito parcial para obtenção do grau de Mestre  
em Ciência da Computação.

ORIENTADOR: Prof. Vinicius Cardoso Garcia

COORIENTADOR: Prof. Thiago Carvalho de Sousa

Aprovada em: 22 / 01 / 2016

Horário: 16 hs

**BANCA EXAMINADORA:**

**Vinicius Cardoso Garcia**

Presidente da Banca

**Kiev Santos da Gama**

1º Examinador

**Vanilson André de Arruda Burégio**

2º Examinador

*A minha mãe Irene Barbosa da Silva, que mesmo com poucos recursos conseguiu me educar,  
moldando em mim o caráter de honestidade e responsabilidade que tenho hoje!*

## **AGRADECIMENTOS**

Esse é um momento em que reconhecemos o quanto nós somos ajudados durante uma caminhada como essa e chegamos à conclusão de que não se consegue nada sozinho. Por isso é muito importante agradecer a todos que nos ajudaram.

Primeiro, agradeço a Deus pelo dom da vida e por ter colocado em minha vida pessoas de bom coração.

Agradeço aos meus familiares, que acreditaram em mim e me deram total apoio nessa empreitada.

Agradeço às minhas filhas, Maria Luiza, Maria Alice e Raissa Lorena, pois elas são minha motivação para lutar com força e perseverança.

Agradeço a todo o corpo docente do Mestrado Profissional do Centro de Informática (CIn) da Universidade Federal de Pernambuco (UFPE), que não mediu esforços para me proporcionar um ensino de qualidade. Em especial, aos meus orientadores, Professor Doutor Vinicius Cardoso Garcia (UFPE) e Professor Doutor Thiago Carvalho de Sousa (UESPI) que despertaram em mim o senso crítico e análise na escrita, aos Professores Ricardo Andrade Lira (UFPI) e a Jefferson Henrique e Lucas Lopes, pela força que me deram na utilização da ferramenta WEKA.

Não poderia deixar de agradecer as empresas participantes que foram peças fundamentais nessa pesquisa.

Agradeço ainda à minha digníssima esposa, Joelma Gualter Silva, pelo apoio moral e espiritual que tanto me deu durante essa caminhada.

Por fim, deixo aqui os meus agradecimentos a todos aqueles que me ajudaram de forma direta ou indireta a alcançar esse objetivo.

*"Não se pode caminhar em direção  
ao futuro com os pés no passado."*

*Marcio Kühne*

## RESUMO

**Contexto:** A Computação em Nuvem apresenta um novo conceito de terceirização na contratação de serviço, esses avanços vêm sendo vistos como uma nova possibilidade para a redução nos volumes dos investimentos em TIC, proporcionados pela maior flexibilidade nos serviços ofertados sob demanda, tendo na redução de custo seu apelo mais forte. Mesmo sabendo dos benefícios do investimento em nuvem, presume-se que algumas empresas são receosas na contratação de serviços e/ou infraestruturas de TIC da computação em nuvem. Essa realidade, apresentada na pesquisa da *Tech Supply*, especializada em Inteligência Tecnológica para Auditoria e Integridade Corporativa e TI, segundo a qual 43% das empresas brasileiras não se sentem seguras para migrar os seus sistemas para nuvem.

**Objetivo:** Nesse contexto geral, apresentam-se dois objetivos: investigar os indícios pelos quais algumas empresas podem estar propensas a contratarem ou não os serviços de Computação em Nuvem no Brasil. Adicionalmente, identificar os motivos, para aquelas que já a usam, de sua satisfação ou insatisfação em relação aos serviços de nuvem contratados no Brasil.

**Método:** Para este estudo, definiu-se o tipo de pesquisa realizada como exploratória de natureza descritiva e explicativa, com ênfase na abordagem quantitativa. Quanto ao procedimento técnico, aplicou-se um levantamento através de um *Survey*, utilizando-se o instrumento de um questionário com 14 (quatorze) itens. Referente à coleta dessas informações, disponibilizou-se através de um formulário WEB (Online). E, por fim, quanto ao tipo de análise aplicada aos resultados, utilizou-se o aprendizado automático para extração dos resultados. Com o uso de aprendizado automático, faz-se necessário o estabelecimento de algumas definições em relação aos métodos de aprendizagem a serem aplicados, como tarefa de classificação por árvore de decisão com algoritmo de classificação J48, método de aprendizagem por indução. Para o modo de treinamento, aplicou-se o não incremental. Na hierarquia do aprendizado, utilizou-se o aprendizado supervisionado e para o paradigma de aprendizado, usou-se o simbólico. Definiram-se também as variáveis classificadoras para cada linha de investigação: “SIM” propensos a contratar ou “NÃO”, para as empresas que não usam; e “SATISFEITO” ou “INSATISFEITO” com a nuvem, para as empresas que já usam.



**Resultado:** Descobriu-se que as características das empresas que estão propensas a contratar a nuvem são garantia de entrega e qualidade dos serviços. Em contrapartida, as empresas que não estão propensas a contratar os serviços da nuvem têm como características o baixo faturamento e poucos colaboradores associados à confiabilidade e segurança da informação. Para a outra linha de investigação, em relação à satisfação, os motivos são o preço da nuvem associado aos modelos de Infraestrutura e Software como Serviço. Por outro lado, para as empresas que estão insatisfeitas, os motivos são segurança da informação, disponibilidade dos serviços associados à redução de custo.

**PALAVRAS-CHAVE:** Computação em Nuvem. Algoritmo Inteligente. Classificação. Extração do conhecimento. Aprendizado automático. *Survey*.

## ABSTRACT

**Context:** Cloud computing presents a new concept of outsourcing at hiring services, these advances have been seen as a new possibility for reduction at volume of investments in ICT, provided for greater flexibility in offered on-demand services, with cost reduction its strongest appeal. Even though the cloud investment benefits, it's assumed that some companies are afraid for contracting services and / or cloud computing ICT infrastructure. This reality presented in the Tech Supply's research, specializing in Technology Intelligence for Audit and Corporate Integrity and IT, according to which 43% of Brazilian companies do not feel safe to migrate their cloud systems.

**Objective:** In general, there are two objectives: to investigate the evidence by which some companies may be prone to hire or not the Computing Cloud services in Brazil. In addition, identify the reasons for those that already use their satisfaction or dissatisfaction with the cloud services contracted in Brazil.

**Method:** For this study, the type of research conducted was defined as exploratory of descriptive and explanatory nature, with an emphasis on quantitative approach. As for the technical procedure, was applied a survey through a Survey, using the instrument of a questionnaire with 14 (fourteen) items. Concerning the collection of this information, it made available through a web form (Online). Finally, the type of analysis applied to the results, we used the automatic learning for extracting results. With the use of automatic learning, it is necessary to establish some definitions regarding learning methods to applied as a classification task by decision tree classification algorithm J48<sup>1</sup>, learning method for induction. For the training mode, applied to the non-incremental. In the learning hierarchy, we used supervised learning and the learning paradigm, was used the symbolic. The classification variables was defined for each research line: "YES" likely to hire or "NO" for companies that do not use; and "SATISFIED" or "DISSATISFIED" with the cloud, for companies that already use.

**Result:** It found that the characteristics of companies that are likely to hire the cloud are delivery assurance and service quality. Conversely, companies that are not likely to hire cloud services characterized by low turnover and few employees associated with the reliability and information security. For another line of research in relation to satisfaction, the reasons are the price associated with cloud models Infrastructure

and Software as a Service. On the other hand, for companies that are dissatisfied, the reasons are information security, availability of services associated with cost reduction.

**Keywords:** Cloud Computing. Intelligent Algorithm. Classification. Knowledge Extraction. Automatic Learning. Survey.

## LISTA DE FIGURAS

FIGURA 1: Árvore de Decisão – Faturamento Entre 2 Milhões e 10 Milhões .....	68
FIGURA 2: Árvore de decisão – Faturamento, Segmento e Motivo específico .....	70
FIGURA 3: Árvore de decisão – Faturamento (Menos que 500 mil), Segmento e Total Colaborador.....	71
FIGURA 4: Árvore de Decisão – Faturamento, Segmento, Motivo Específico e Tipo de Nuvem .....	72
FIGURA 5: Árvore de decisão – Faturamento e Segmento .....	73
FIGURA 6: Árvore de decisão – Outros faturamentos .....	75
FIGURA 7: Árvore de decisão- Faturamento Segmento e Características .....	76
FIGURA 8: Árvore de decisão – Faturamento Segmento e Tipo de Serviço.....	77
FIGURA 9: Árvore de Decisão – Faturamento, Segmento, Características e Redução de custo.....	79
FIGURA 10: Árvore de decisão – Faturamento, Segurança da Nuvem Pública e Total Colaborador.....	80

## LISTA DE GRÁFICOS

GRÁFICO 1: Resultado Geral .....	54
GRÁFICO 2: Área de Atuação .....	56
GRÁFICO 3: Faturamento anual .....	56
GRÁFICO 4: Quantidade de colaboradores .....	57
GRÁFICO 5: Tipo de nuvem .....	58
GRÁFICO 6: Motivos específicos .....	59
GRÁFICO 7: Tipos de serviços .....	60
GRÁFICO 8: Provedor de serviço .....	61
GRÁFICO 9: Serviço utilizado .....	62
GRÁFICO 10: Características do serviço .....	63
GRÁFICO 11: Tipo de Nuvem .....	64
GRÁFICO 12: Segurança da Nuvem Pública .....	65
GRÁFICO 13: Redução de custos .....	66

## LISTA DE TABELAS

TABELA 1: Matriz de Confusão – Treinamento Contrataria ou não nuvem .....	73
TABELA 2: Matriz de Confusão – Treinamento Satisfeita ou Insatisfeita .....	81
TABELA 3: Casos classificados com 100% de acertos.....	91
TABELA 4: Casos classificados com Acertos/Erros (Contrataria ou Não Nuvem) ....	92
TABELA 5: Casos com 100% de acertos (Satisfação ou Insatisfação).....	93
TABELA 6: Casos classificados com Acertos/Erros (Satisfação ou Insatisfação).....	94

## LISTA DE ABREVIATURAS

ARPANET.....	Advanced Research Projects Agency Network
API's.....	Application Programming Interface
CRM.....	Customer Relationship Management
CIN.....	Centro de Informática.
CIO.....	Chief Information Officer
CEO.....	Chief Executive Officer
CRISP-DM.....	Cross Industry Standard Process for Data Mining.
DaaS.....	Banco de Dados como Serviço.
DTI's.....	Diretores de Tecnologia da Informação.
EC2.....	Elastic Computação em Nuvem.
ERP.....	Enterprise Resource Planning.
FaaS.....	Perícia Digital como Serviço.
IaaS.....	Infraestrutura como Serviço.
IDC.....	Internet Data Center.
KDD.....	Knowledge Discovery in Databases.
NIST.....	National Institute of Standards and Technology.
PaaS.....	Plataforma como Serviço.
PC.....	Personal Computer.
QoS.....	Quality of Service.
ROI.....	Retorno sobre Investimento.
RBT.....	Risk-Based Testing.
Las.....	Service Level Agreements.
SaaS.....	Software como Serviço.
SOAP.....	Simple Object Access Protocol.
SPSS.....	Statistical Package for the Social Sciences.
STaaS.....	Armazenamento como Serviços.
TIC.....	Tecnologia da Informação e Comunicação.
TI.....	Tecnologia da Informação.
TCO.....	Custo Total de Propriedade.
UFPE.....	Universidade Federal de Pernambuco.
UFPI.....	Universidade Federal do Piauí.
UESPI.....	Universidade Estadual do Piauí.
WDSL.....	Web Services Description Language.

## SUMÁRIO

<b>1 INTRODUÇÃO .....</b>	<b>18</b>
1.1 Estabelecimento do Problema .....	19
1.2 Justificativas .....	20
1.3 Resultados esperados .....	20
1.4 Metodologia.....	21
1.5 Escopo negativo.....	22
1.6 Estrutura da Dissertação.....	23
<b>2 REFERENCIAL TEÓRICO .....</b>	<b>23</b>
2.1 Conceitos e definições .....	24
2.2 Tecnologias habilitadoras da Nuvem.....	26
2.2.1 Computação utilitária .....	26
2.2.2 Clusterização .....	27
2.2.3 Computação em Grade.....	27
2.2.4 Virtualização .....	28
2.3 Características.....	29
2.3.1 Tipologia da Nuvem .....	31
2.3.2 Modelos de serviços oferecidos.....	31
2.4 Vantagens e Desvantagens .....	32
2.4.1 Vantagens .....	32
2.4.2 Desvantagens.....	33
2.5 Aprendizado automático.....	34
2.5.1 Classificação por Árvore de Decisão .....	35
2.6 Adoção/Migração Para Computação em Nuvem .....	36
2.7 Desafios para essa migração no Brasil .....	37
2.8 Resumo do capítulo.....	40
<b>3 METODOLOGIA DA PESQUISA .....</b>	<b>41</b>
3.1 Definição metodológica .....	41
3.2 Métodos do aprendizado automático .....	43
3.3 Propósitos da Pesquisa.....	45
3.4 Planejamento Amostral (Questionamentos).....	45
3.4.1 Perspectivas da identificação da Empresa ou Organização (I):.....	46
3.4.2 Perspectivas das empresas que ainda não aderiram à nuvem (II): .....	46



3.4.3 Perspectivas das empresas que fazem uso da tecnologia nuvem (III):	47
3.5 Público-alvo, Pré-testes e Meios de divulgação	48
3.6 Recursos utilizados (Ferramentas)	49
3.7 Resumo do Capítulo	51
<b>4 COLETA DOS DADOS</b>	<b>52</b>
4.1 Dados preliminares	53
4.1.1 Segmento de atuação	53
4.1.2 Quanto ao faturamento anual	54
4.1.3 Quadro efetivo (Quantidade de colaboradores)	55
4.1.4 Responsável pela TIC	56
4.1.5 Quanto à utilização da Nuvem	56
4.2 Sob a perspectiva das Empresas ou Organizações que NÃO utilizam Computação em Nuvem	57
4.2.1 Quanto ao tipo de nuvem	57
4.2.2 Motivos específicos	58
4.2.4 Quanto à migração para um Serviço	58
4.2.5 Quanto ao provedor do Serviço	59
4.3 Sob a perspectiva das Empresas ou Organizações que utilizam Computação em Nuvem	60
4.3.1 Quanto aos serviços utilizados	60
4.3.2 Quanto às características (Serviços utilizados)	61
4.3.3 Quanto ao tipo de nuvem utilizada	62
4.3.4 Segurança da Nuvem	63
4.3.5 Quanto à redução de custo	64
4.4 Resumo do Capítulo	65
<b>5 ANÁLISE DOS DADOS</b>	<b>66</b>
5.1 Treinamentos	66
5.1.1 Treinamento para identificar os indícios para contratar ou não	67
5.1.2 Treinamento para identificar o motivo da Satisfação ou Insatisfação	73
5.1.2.1 Matriz de Confusão – Treinamento Satisfeita ou Insatisfeita	80
5.2 Resumo do Capítulo	81
<b>6 CONSIDERAÇÕES FINAIS</b>	<b>82</b>
6.1 Ameaças aos resultados	82
6.2 Trabalhos correlatos e trabalhos futuros	83

6.2.1 Trabalhos correlatos .....	83
6.2.2 Trabalhos Futuros .....	88
6.3 Descobertas da pesquisa .....	89
6.3.1 Hipótese de contratar ou não .....	90
6.3.2 Hipóteses de Satisfação ou Insatisfação .....	92
6.4 Conclusões .....	94
<b>REFERÊNCIAS .....</b>	<b>96</b>
<b>APÊNDICES .....</b>	<b>102</b>

## 1 INTRODUÇÃO

A Tecnologia da Computação em Nuvem pode ser considerada atualmente como uma tecnologia inovadora que introduz novas ideias de negócios nas organizações (BAUN *et al.*, 2011). Para Linthicum (2010), a Computação em Nuvem é um movimento que está promovendo uma disruptura nos departamentos de TIC. Esse movimento faz com que quanto maior a quantidade de serviços de nuvem contratados, mais os modelos tradicionais de fornecimentos de hardware e software corporativo são encolhidos (LINTHICUM, 2013). A tecnologia Nuvem está se disseminando e transformando as empresas e até mesmo aquelas corporações que se acham distantes dessas transformações que serão afetadas, mais cedo ou mais tarde (TAURION, 2015).

Aliada a essa transformação e disseminação há o novo paradigma de arquitetura da TIC, a Computação em Nuvem, tema central deste trabalho, que está potencializando o crescimento dos negócios em virtude de características singulares, tais como: redução de custo, escalabilidade, elasticidade, cobrança de acordo com o consumo e serviço sob demanda (MELL; GRANCE, 2009).

Diante de um ambiente dinâmico como a Tecnologia da Informação, a sobrevivência de uma empresa pode depender da sua necessidade de apresentar seus resultados e na sua capacidade de adaptar-se às mudanças. Nesse contexto, surge a Computação em Nuvem, que permite criar servidores virtuais em um curto espaço de tempo.

Um diferencial e um dos maiores atrativos da Computação em Nuvem é o pagamento de acordo com o consumo, que se assemelha ao fornecimento dos serviços básicos, como de Água, Luz e Telefone (VECCHIOLA *et al.*, 2009). Para esses fornecimentos, não é necessário saber suas origens, bastando simplesmente utilizá-lo.

Outro fator interessante é que, em um sistema de Computação em Nuvem, os computadores locais já não têm mais que fazer todo o trabalho pesado de uma empresa, quando se trata de executar aplicações. A demanda por hardware e software por parte do usuário diminui bastante, não sendo mais necessárias a instalação ou configuração dos sistemas em ambientes locais. O usuário precisa ser capaz de executar o software, por meio da interface do sistema na Computação em

Nuvem, que pode ser tão simples como a utilização de um navegador *Web* (MELL; GRANCE, 2009).

## 1.1 Estabelecimento do Problema

No atual cenário, os recentes avanços na disponibilização de soluções de Computação em Nuvem vêm sendo vistos como uma nova possibilidade para a redução nos volumes dos investimentos em TIC, proporcionando maior flexibilidade nos serviços ofertados sob demanda, tendo na redução de custo seu apelo mais forte (BUBLITZ, 2010). Entretanto, mesmo sabendo dos benefícios do investimento em nuvem, presume-se que algumas empresas são receosas na contratação de serviços e/ou infraestruturas de TIC da computação em nuvem.

Uma pesquisa realizada em 2014 pela Tech Supply<sup>1</sup> revelou que, apesar de todas as facilidades proporcionadas pela nuvem, 43% das empresas brasileiras não se sentem seguras para migrar os seus sistemas para a computação em nuvem. Diante disso, pode-se questionar: por que algumas empresas estão propensas a contratarem ou não os serviços da computação em nuvem? E as que adotaram essa arquitetura em nuvem já têm condições de se afirmarem satisfeitas ou não e por quais motivos?

No contexto geral, este trabalho possui dois objetivos: o primeiro é investigar os indícios pelos quais algumas empresas podem estar propensas à contratação ou não dos serviços de Computação em Nuvem no Brasil. Adicionalmente, identificar os motivos, para aquelas que já a usam, de sua satisfação ou insatisfação em relação aos serviços de nuvem contratados no Brasil.

Como delimitação dos objetivos deste trabalho, define-se os seguintes objetivos específicos:

- a) Evidenciar os indícios para contratar ou não;
- b) Identificar as variáveis influenciáveis na decisão de contratar ou não;
- c) Coletar amostra a partir das informações dos respondentes;
- d) Analisar os dados da coleta através das amostras apresentadas;
- e) Identificar os motivos de satisfação ou não das empresas que já utilizam os serviços em nuvem.

---

<sup>1</sup> Empresa especializada em inteligência Tecnológica para Auditoria e Integridade Corporativa de TI. < <http://www.techsupply.com.br/>> Acesso em 23 de janeiro de 2016.

<sup>2</sup> Algoritmo de aprendizado para extração de conhecimento recodificado na linguagem JAVA, a partir

- f) Descobrir conhecimento através do aprendizado automático.
- g) Apresentar hipótese conclusiva em relação ao aprendizado realizado.

## **1.2 Justificativas**

Pode-se justificar a realização deste estudo empírico e investigativo na utilização dos resultados que podem ser aproveitados tanto pelos pesquisadores, como pelas empresas, sem se restringindo ao ambiente acadêmico. Outras justificativas são:

- Fornecer resultados que ajudem a compreender melhor: os indícios da não adesão; e o nível de satisfação das empresas que estão utilizando nuvem.
- Especificamente, a investigação dos indícios da “não migração” e satisfação ou insatisfação.
- Aplicar o aprendizado automático nos dados coletados com o intuito de apresentar resultados mais precisos.

## **1.3 Resultados esperados**

Espera-se com esta pesquisa apresentar resultados, por meio do aprendizado automático, que auxiliem as empresas em seus processos migratórios para esse ambiente da Computação em Nuvem, podendo facilitar o entendimento das variáveis que norteiam a adesão. Para corroboração dessa primeira hipótese, aposta-se na identificação dos motivos para empresas contratar ou não algum serviço da nuvem.

Em relação à outra expectativa de investigação, o resultado esperado é a identificação dos níveis de satisfação ou insatisfação por parte das empresas brasileiras que se encontram utilizando algum tipo de serviço da Computação em Nuvem. Para tanto, sustenta-se a hipótese da satisfação, principalmente dos contratantes, tendo como parâmetro os benefícios que a Nuvem oferece. Espera-se também a boa aceitação desses resultados por parte da comunidade acadêmica e profissional.

## 1.4 Metodologia

A pesquisa realizada baseia-se em um estudo criterioso sobre a Computação em Nuvem a fim de apresentar um quadro representativo da realidade brasileira em relação à migração e satisfação, tendo o sistema tradicional como objeto de referência. Mais especificamente, a pesquisa buscou informações sob os métodos apresentados no viés da metodologia científica.

Para este estudo, define-se o tipo de pesquisa a ser realizada como exploratória de natureza descritiva e explicativa, com ênfase na abordagem quantitativa, que possibilita o uso de dados estatísticos para a interpretação dos resultados coletados. Como procedimento técnico, apoia-se em levantamento através de um *Survey*, utilizando como instrumento um questionário com 14 (quatorze) itens. A coleta dessas informações se deu a partir de um formulário *WEB (online)*. E, por fim, o tipo de análise aplicada aos resultados é o aprendizado automático para extração dos resultados.

Com o uso de aprendizado automático, faz-se necessário o estabelecimento de algumas definições em relação aos métodos de aprendizagem a serem aplicados, principalmente visando os melhores resultados do treinamento. Para o aprendizado, utilizou-se tarefa de classificação por árvore de decisão com algoritmo de classificação J48<sup>2</sup> (ZUBEN; ATTUX, 2010). Quanto ao método de aprendizagem, utilizou-se aprendizado por indução, que permite aprender a partir de conjunto de exemplos. (CHAPELLE *et al.*, 2006).

Após definir o método, é necessário definir o modo de treinamento. Aplicou-se o método não incremental, pois não permite acrescentar um novo exemplo a partir do conjunto de treinamento atual (STANGE; NETO, 2011). Na hierarquia do aprendizado, utilizou-se o aprendizado supervisionado, pois possui a figura representativa do “Professor” como medida de classificação dos dados (LUGER, 2007). Em relação ao paradigma de aprendizado, aplicou-se o simbólico, que possui no aprendizado um conjunto de representação simbólica (DE PAULA, 2002). Quanto à configuração dos treinamentos, aplicou-se a configuração “com poda”, que consiste em classificar os casos mais significativos do treinamento (ZUBEN; ATTUX,

---

<sup>2</sup> Algoritmo de aprendizado para extração de conhecimento recodificado na linguagem JAVA, a partir do algoritmo C4.5 (Escrito por Ross Quinlan, em 1993), que foi escrito na linguagem C (REVINT, Nº 1 2013).

2010). E, por fim, definiu-se as variáveis classificadoras de acordo com a abordagem da pesquisa: “SIM”, quando propensas a contratar, ou “NÃO”, para a outra linha de investigação; o classificador “SATISFEITO” com a utilização da nuvem e “INSATISFEITOS” com a nuvem.

### **1.5 Escopo negativo**

Algumas variáveis não foram abordadas nesta pesquisa, embora estejam de algum modo relacionadas com o tema da Computação em Nuvem. Por exemplo, não foi questionada a existência de profissionais capacitados com o objetivo de suprir essa mão de obra especializada no mercado brasileiro, abordando se o mercado de trabalho brasileiro está preparado para suprir a demanda da Computação em Nuvem com mão de obra qualificada. Para as empresas que pretendem adquirir os serviços da Nuvem, é importante ter conhecimento dessa informação, uma vez que, devido a ela, a empresa pode se sentir mais confiante no momento de migrar.

Outra variável que não foi tratada nessa pesquisa como critério de contratação ou não da nuvem, foi a portabilidade da Nuvem. Sabe-se que a flexibilidade com que se apresenta a computação em Nuvem faz dela uma tecnologia com bastante potencial e, embora os ambientes em nuvem sejam flexíveis, o aspecto da portabilidade deve ser observado pelos contratantes no momento de migrar seus negócios e/ou infraestruturas, portabilidade tanto de aplicativos como de dados. Para quaisquer modelos de serviços em Nuvem, ela pode contribuir para deixar o modelo mais dinâmico e flexível. É possível que as empresas mudem de provedor de serviço por algum motivo particular, para isso ambos os lados (origem e destino) devem estar preparados para suportar esse impacto de mudança.

Diante disso, as empresas devem possuir um modelo com a capacidade de adaptação e possibilitar a mudança para outro provedor de serviço sem que haja os desconfortos dos recursos do ambiente (como por exemplo, incompatibilidade entre plataforma, dados ou aplicativos). Na visão da Skyinone (2014, p. 4), “o uso da nuvem é irreversível, é muito raro, senão impossível, vermos uma regressão tecnológica”, uma vez que uma organização constrói ou migra de um sistema para usar em Nuvem, trazer esse sistema de volta para o ambiente anterior ou até mesmo migrar para um novo provedor Nuvem pode se tornar difícil e demanda muito

recurso financeiro. A portabilidade “é uma medida utilizada para determinar o impacto do movimento do consumidor de nuvem entre os recursos de TI e nuvens” (ERL, 2005, p. 47) e é também “a capacidade de migrar uma aplicação de um serviço de nuvem para outro” (CARRION *et al.* 2014, p. 1). Certamente, essas são informações (profissionais especializados e portabilidade da nuvem) melhorariam os resultados dessas linhas de investigação, mas em virtude da delimitação do escopo, não fazem parte dessa abordagem.

## 1.6 Estrutura da Dissertação

Esta dissertação conta com esta introdução, seguida de mais cinco capítulos:

- **Capítulo 2 (Referencial Teórico):** realiza-se a fundamentação teórica, abordando os acontecimentos que antecederam o paradigma da Computação em Nuvem, desde o surgimento do termo até seus conceitos mais recentes, como classificação e categorização de serviços e tipos de nuvem ofertados.
- **Capítulo 3 (Metodologia da Pesquisa):** abordam-se os conceitos, tais como: modelo da pesquisa, propósito da pesquisa, tópicos da investigação, planejamento amostral, público-alvo, pré-testes, meios de divulgação e recursos utilizados (ferramental).
- **Capítulo 4 (Coleta dos Dados):** geram-se as frequências dos respondentes através dos gráficos e representações dos dados, utilizando para isso as ferramentas de estatística apropriadas.
- **Capítulo 5 (Análise dos Dados):** realiza-se a análise dos dados, extraíndo conhecimento através do aprendizado automático.
- **Capítulo 6 (Considerações Finais):** apresentam-se as considerações finais, os trabalhos correlatos e trabalhos futuros, seguidos das descobertas da pesquisa e suas conclusões. Por fim, apresentam-se os Apêndices A e B, que contêm os formulários desenvolvidos especificamente para viabilização da pesquisa a partir da ferramenta *LimeSurvey*<sup>3</sup> e os treinamentos realizados na ferramenta WEKA<sup>4</sup> para o aprendizado automático do conjunto de treinamento.

---

<sup>3</sup> Ferramenta de desenvolvimento de pesquisas online. Disponível em: <http://www.limesurvey.com> Acesso em: 16 de novembro de 2015.

<sup>4</sup> *Waikato Environment Knowledge Analysis – WEKA*. Disponível em: <https://www.weka-akademie.de/?cholid=00898157> Acesso em: 15 de novembro de 2015.



## 2 REFERENCIAL TEÓRICO

Este capítulo tem como objetivo explanar os conceitos e estudos utilizados como base para a presente pesquisa.

Na seção 2.1 encontram-se os conceitos e definições do tema; na seção 2.2 são apresentadas as tecnologias habilitadoras da Computação em Nuvem; a seção 2.3 apresenta as principais características da Computação em Nuvem, como *selfservice* sob demanda, amplo acesso à rede, *pooling* de recursos, elasticidade rápida, tipologia de Nuvem e modelos de serviços ofertados; a seção 2.4 traz as vantagens e desvantagens da Nuvem; na seção 2.5 apresentam-se os conceitos de aprendizado automático e tarefa de classificação por árvore de decisão; a seção 2.6 aborda a adoção/migração para Computação em Nuvem; a seção 2.7 trata dos desafios para essa migração no Brasil, considerando as leis e seus impactos; e, por fim, a seção 2.8 apresenta o resumo do capítulo.

### 2.1 Conceitos e definições

“A Nuvem está em toda parte. Abra qualquer revista de tecnologia ou visite qualquer site ou blog e certamente você verá alguma matéria sobre Nuvem” (VELTE *et al.* 2010, p. 3). Mas nem todo mundo sabe exatamente o que é Computação em Nuvem. Para entender melhor a estrutura e o funcionamento da Computação em Nuvem, deve-se começar pelos conceitos abordando os acontecimentos que possibilitaram o surgimento da Nuvem.

Inicialmente, os dispositivos eram basicamente computadores de mesa, estações de trabalho e os chamados servidores, que armazenavam e transmitiam informações, como páginas da WEB e mensagens de e-mail (KUROSE, 2010). A TIC utilizava-se do recurso de compartilhamento da informação como uma de suas principais características e a internet é um dos principais meios para tal.

“A fusão dos computadores e das comunicações promoveu uma profunda influência na forma como os sistemas computacionais são organizados”. (TANENBAUM, 2010, p. 18). O novo paradigma de tecnologia, chamado Computação em Nuvem (do inglês *Cloud Computing*), representa muito bem esse novo cenário gerado. Alguns autores, como Taurion (2015), definem Nuvem como

sendo uma plataforma de fornecimento de serviços, infraestrutura, plataforma e software sob demanda.

Para a AWS<sup>5</sup>, "Computação em Nuvem", por definição, diz respeito à entrega sob demanda de recursos da TIC e aplicativos pela internet, com modelo de definição de preço conforme a utilização. Esses recursos podem ser dinamicamente reconfigurados para se ajustar a uma carga variável (escala), permitindo também uma melhor utilização dos recursos. Esse conjunto de recursos é tipicamente explorado por um modelo *pay-per-use* (pague pelo uso), em que as garantias são oferecidas pelo provedor de infraestrutura por meio de SLAs<sup>6</sup> personalizados. Para Alberto (2010), a SLA é um acordo entre provedor de serviço e seu cliente, determinando a qualidade mínima de serviço aceitável, existindo uma provisão dinâmica de recursos e contratação de diversos serviços da Nuvem.

De acordo com entrevista concedida a Freire (2011), Taurion diz que a Computação em Nuvem não pode ser vista como uma iniciativa de simples cunho tecnológico, de melhoria de infraestrutura. Dada a significância das possibilidades de serviços, a Computação em Nuvem precisa ser considerada sob uma perspectiva mais ampla de estratégia, ou seja, representa muito mais. A sua utilização promove progresso e melhorias qualitativas de serviços, além de provocar um reposicionamento da própria empresa no mercado.

Vale ressaltar que a definição do NIST (*National Institute of Standards and Technology*) é a que melhor representa conceito da Computação em Nuvem e que, por isso, é adotada neste trabalho:

A Computação em Nuvem é um modelo que possibilita acesso, de modo conveniente e sob demanda, a um conjunto de recursos computacionais configuráveis (por exemplo, redes, servidores, armazenamento, aplicações e serviços) que podem ser rapidamente adquiridos e liberados com mínimo esforço gerencial ou interação com o provedor de serviços (MELL; GRANCE, 2009, p. 2, tradução nossa).<sup>7</sup>

---

<sup>5</sup> Empresa americana provedora de serviços em nuvem. *Amazon Web Services*. Disponível em: <http://www.aws.amazon.com>. Acesso em: 16 de agosto de 2015.

<sup>6</sup> Um Acordo de Nível de Serviço (SLA) é uma obrigação contratual entre contratante e seu fornecedor (HURWITZ et al, 2010).

<sup>7</sup> "Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models."

Diante do entendimento dos conceitos, é importante destacar as tecnologias habilitadoras da computação em nuvem, detalhadas no tópico 2.2 a seguir.

## **2.2 Tecnologias habilitadoras da Nuvem**

Para o surgimento da computação em nuvem é importante ressaltar a contribuição de algumas tecnologias precursoras que possibilitaram o aparecimento da Nuvem, são elas: Computação Utilitária, Clusterização, Computação em Grade e virtualização. Cada uma delas é discutida nas subseções a seguir.

### **2.2.1 Computação utilitária**

De acordo com Rittinghouse e Ransome (2009), os conceitos iniciais da Computação Utilitária surgiram no ano de 1961, quando o professor John McCarthy sugeriu que a tecnologia poderia levar a um futuro no qual o poder de computação e até mesmo de aplicações específicas poderiam ser vendidos a partir de um modelo de negócio do tipo utilitário. Essa ideia se tornou muito popular nos anos de 1960. No entanto, desde a virada do milênio, o conceito foi reformulado com o surgimento de novas tecnologias no campo da virtualização.

A origem do termo Computação Utilitária se deve à mudança constante da computação. Nos tempos atuais, as empresas não têm mais a necessidade de adquirir computadores robustos (supercomputadores), tal como era feito há pouco tempo. Hoje o modelo de Computação Utilitária oferece uma série de benefícios para os prestadores de serviços e usuários (MOHAMED, 2009).

De acordo com Mohamed (2009), a Computação Utilitária pode oferecer, por meio de seus servidores, recursos computacionais no formato de mercadoria com um baixo custo. Assim, a Computação em Nuvem “herdou”, de certa forma, as principais características da computação utilitária: seus usuários não necessitam ter uma atenção especial para sua elasticidade, poder de processamento, escalabilidade e capacidade de armazenamento. “Esse modelo foi desenvolvido para suportar sistemas escaláveis” (BORGES *et al.*, 2011, p. 23). A essência do modelo de Computação Utilitária são os recursos de capacidade de armazenamento, tempo de processamento e largura de banda, que são encapsulados como serviços e disponibilizados aos contratantes. Outra tecnologia habilitadora da Nuvem é vista a

seguir, que é a Clusterização. Ela permite trabalhar ambiente de processamento da informação com estrutura homogênea.

### 2.2.2 Clusterização

“A idéia inicial deste sistema surgiu em 1960, quando a IBM, interligou os seus mainframes e obteve ganho no processamento através do paralelismo” (CONTI, 2009, p. 3). A Clusterização possui aspectos que reúne dois ou mais computadores para que trabalhem de maneira conjunta com o objetivo de realizar tarefas. Porém, para compartilhar esses recursos, o sistema clusterizado possui alguns requisitos obrigatórios, como por exemplo, ter arquiteturas iguais e possuir uma estrutura homogênea. “Para reforçar estes aspectos, pode-se até mesmo, adotar sistemas operacionais preparados especialmente para *clustering*” (ALECRIM, 2013, p.7).

De acordo com Buyya (1999), Cluster é um tipo de sistema para processamento paralelo e distribuído que consiste numa coleção de computadores interconectados e trabalhando juntos como um único recurso computacional integrado. As soluções desenvolvidas em arquitetura *cluster computing* é uma solução para disponibilizar alta performance e alta escalabilidade dentro de uma infraestrutura própria e homogênea, sem a necessidade de interligar os *nós* da rede em locais geograficamente distantes (COLVERO *et al*, 2005). Um ambiente clusterizado geralmente é desenvolvido para proporcionar recursos computacionais com alto poder de processamento, performance e desempenho.

Na sequência, temos a tecnologia da Computação em Grade, que possui características semelhantes à Clusterização.

### 2.2.3 Computação em Grade

“Esse sistema computacional surgiu nos anos 90, para suprir com a demanda por alto desempenho computacional” (CONTI, 2009, p. 4). O termo Computação em Grade é inspirado no conceito de *Power Grid*, que foi idealizado por Ivan Foster no ano de 1998. Esse termo designa o sistema de geração e distribuição de energia. Também conhecida como Computação em Grade, o seu funcionamento é baseado na utilização numa rede (FOSTER; KESSELMAN, 2003).

Os participantes dessa arquitetura conectam-se através de sistemas distribuídos com um gerenciamento dos recursos de hardware, rede, conexão com a banda larga e processamento dos dados. A Computação em Grade possui algumas características semelhante à Computação em Cluster, tais como, escalabilidade, conjunto de computadores interligados, compartilhamento de recursos, estrutura homogênea, pervasividade (plugar e usar), visão computacional igual, armazenamento remoto de dados e distribuição em larga escala. Se comparadas às características existentes na Computação em Nuvem, algumas características se apresentam iguais.

A seguir, temos a Virtualização, mais uma tecnologia habilitadora da computação em Nuvem.

#### 2.2.4 Virtualização

Há dez anos, quando a Computação em Nuvem ainda estava sendo apresentada, seu foco era serviços simples, dentro de uma infraestrutura pública. Assim como outros modelos, o que é bastante comum na tecnologia, ela evoluiu na sua forma de uso. A virtualização passou por processo semelhante, desde a virtualização de hardware e seus sistemas operacionais, até a virtualização de redes e serviços.

A ideia de virtualização surgiu na década de 1960, a partir de um projeto idealizado na *International Business Machines* (IBM), com o objetivo maior de melhorar a utilização dos recursos de um servidor *mainframes* (máquinas e servidores de grande porte). O surgimento e crescimento de outros modelos computacionais e, principalmente, da arquitetura cliente/servidor fez com que o projeto idealizado fosse descontinuado. Após três décadas, o conceito de virtualização voltou com mais entusiasmo.

O ambiente virtual possui uma ideia de computação distribuída, virtualizando máquinas, tecnologias e métodos para dar suporte a uma estrutura de *softwares* (BAUN *et al.*, 2011). Nos últimos anos, com a evolução da tecnologia, o poder de processamento dos computadores aumentou significativamente, no entanto, em geral, todo esse poder de processamento estava sendo subutilizado.

Uma das qualidades da virtualização é dispor desses recursos computacionais e seu objetivo é oferecer esses recursos a quem deles necessita

para trabalhar. A virtualização pode ser implantada de duas formas: na forma total e, para este modelo, é construída uma réplica da máquina física, de modo que não há necessidade de que o sistema operacional visitante seja modificado para ter ciência de que está sendo executado em uma máquina virtual; e na forma de paravirtualização, em que o sistema operacional visitante tem que ser modificado para ser executada a máquina virtual. Entre as vantagens da virtualização, destacam-se a melhor exploração dos recursos de processamento, a economia energética e a redução no espaço físico utilizado.

Diante do exposto, é possível afirmar que a Computação em Nuvem pode ser considerada o termo sucessor da virtualização, pois tem entre suas características a virtualização de infraestrutura, plataforma, software e outros serviços (CARISSIMI, 2008).

## 2.3 Características

Para Mahmood (2011), a Computação em Nuvem está se tornando uma das palavras-chave da indústria de TIC. A nuvem é uma metáfora para a Internet ou infraestrutura de comunicação entre os componentes arquiteturais, baseada em uma abstração que oculta a complexidade de infraestrutura. Cada parte dessa infraestrutura é provida como um serviço e esses serviços são, normalmente, alocados em *data centers*, utilizando hardware compartilhado para computação e armazenamento (SOUSA, 2012).

O NIST apresentou em 2009 um documento que é referência para examinar as características essenciais que a Computação em Nuvem possui. São elas: *selfservice* sob demanda, amplo acesso à rede, *pooling* de recursos, elasticidade rápida e serviço medido. Também são apresentados os modelos de serviço, juntamente com seus tipos de implantação da Nuvem. Essas características representam o modelo de Computação em Nuvem para este trabalho (MELL; GRANCE, 2009 apud SOLANGE, 2015), conforme o que segue:

- **Alocação de recursos sob demanda:** permite que o usuário possa dimensionar a infraestrutura necessária de recursos computacionais sob demanda. Similar a um serviço sob demanda, esta característica permite que usuários solicitem recursos em tempo de execução à medida que necessitar;

- **Ampla acesso à rede:** os recursos são disponibilizados através do ambiente de rede e devem estar disponíveis para acesso por meio de uma gama de dispositivos como *tablets*, PCs, *smartphones*, entre outros.

- **Pooling de Recursos:** os recursos computacionais do provedor de serviço são estruturados para servir a múltiplos usuários utilizando um modelo *multi-tenant* (MT), que disponibiliza diferentes recursos físicos e virtuais de maneira dinâmica, conforme a necessidade do usuário. Há um senso de independência local, ou seja, o usuário não precisa ter conhecimento da localização física dos recursos computacionais, bastando apenas especificar a localização em um nível de abstração mais alto (país, estado etc.).

- **Elasticidade e Escalabilidade:** a elasticidade é a característica que permite que os recursos disponíveis ao usuário pareçam ilimitados, pois tais recursos podem ser adicionados e removidos de maneira rápida e automática, conforme a necessidade da carga de trabalho. Por seu turno, a escalabilidade está relacionada com o requisito de aumento da capacidade de trabalho mediante a adição proporcional de recursos.

- **Serviço medido:** os recursos de um provedor Nuvem são automaticamente controlados e otimizados pela capacidade de medição em um nível de abstração adequado para o tipo de serviço. A utilização dos recursos pode ser controlada, monitorada e relatada com transparência entre o provedor e consumidor do serviço. Borges *et al.* (2011) entendem que esse monitoramento agrega transparência tanto para o provedor quanto para o cliente, sendo que normalmente são utilizados contratos referentes aos serviços (SLA) para especificar as características dos serviços, parâmetros de qualidade (QoS) e determinar os valores que serão cobrados. Um SLA define os níveis de disponibilidade, funcionalidade, desempenho e outros atributos relativos aos serviços, incluindo penalidades para o caso de violação das regras por qualquer uma das partes.

- **Disponibilidade:** os recursos devem estar disponíveis através da rede, estando acessíveis por meio de dispositivos computacionais padrões, promovendo sua utilização por plataformas heterogêneas, como telefones celulares, *laptops*, PDAs etc. Dessa forma, a Nuvem, aparentemente, seria um ponto de acesso centralizado para as necessidades computacionais dos seus usuários, estando disponível o tempo todo e em qualquer lugar (BORGES *et al.*, 2011).

Além das características apresentadas pelo NIST, destacam-se também a tipologia da nuvem e os modelos de serviços oferecidos. Cada uma dessas características é tratada com mais atenção nas subseções a seguir.

### 2.3.1 Tipologia da Nuvem

Antes mesmo de se tratar do modelo de implantação da Computação em Nuvem, é importante ressaltar que os serviços oferecidos em nuvem possuem uma forte tendência a serem utilizados em diversas empresas, independentemente de seu porte. Em relação à forma como a Computação em Nuvem é disponibilizada ao usuário final, Mell e Grance (2009) apresentam as seguintes categorias:

- **Nuvem privada:** a infraestrutura de Nuvem é provisionada para uso exclusivo de uma única organização compreendendo vários consumidores (por exemplo, unidades de negócio);
- **Nuvem comunitária:** o ambiente de Nuvem é provisionado para uso exclusivo por uma determinada comunidade de consumidores compartilhando interesses comuns (por exemplo, requisitos de segurança, monitoramento e outros);
- **Nuvem pública:** o ambiente da Nuvem é provisionado para uso aberto pelo público em geral. Pode ser gerenciado e operado por uma empresa, meio acadêmico ou organizacional;
- **Nuvem híbrida:** é uma composição de dois ou mais tipos de Nuvens distintas quanto à sua infraestrutura (privada, comunitária ou pública), que permanecem entidades únicas.

### 2.3.2 Modelos de serviços oferecidos

No ambiente da Computação em Nuvem, há vários modelos de serviços. Esses ambientes permitem que sejam definidos padrões de arquitetura de acordo com as necessidades do momento. Quanto aos tipos de serviços oferecidos em Computação em Nuvem, Erl (2005) apresenta as seguintes classificações:

- **Software as a Service (SaaS):** os aplicativos são hospedados e entregues *online* através de um navegador web que oferece funcionalidade tradicional, como o Google Docs, Gmail e MySAP;



- **Plataforma como Serviço (PaaS):** fornece ao consumidor um ambiente onde ele possa desenvolver utilizando linguagens de programação e ferramentas suportadas pelo fornecedor;
- **Infraestrutura como Serviço (IaaS):** é um ambiente que possui um conjunto de recursos de computação virtualizados e hospedados em nuvem;
- Outras classificações como **Armazenamento como Serviço (SaaS):** que permite o armazenamento de informações como forma de serviços;
- E **Banco de dados como Serviço – DaaS:** permite o funcionamento do Sistema de Gerenciamento de Banco de Dados (SGDB) a partir do ambiente em nuvem.

## 2.4 Vantagens e Desvantagens

A Computação em Nuvem se apresenta com vários benefícios, dentre eles escalabilidade e pagamento somente do que foi consumido. Mas sabe-se, como toda tecnologia, que também existem desvantagens. Essas duas questões são abordadas nos tópicos seguintes.

### 2.4.1 Vantagens

Uma das vantagens mais proeminentes da Computação em Nuvem é a comodidade, permitindo o acesso aos serviços de qualquer lugar, bastando apenas ter um dispositivo conectado à internet, de preferência com uma banda larga de qualidade, o que proporciona comodidade e flexibilidade aos usuários.

Além dessa, para Taurion (2009), sob uma visão mercadológica, existem algumas outras vantagens da Computação em Nuvem a serem destacadas:

- A administração da nuvem pode se situar em torno de 1/5 do que seria necessário em sistemas distribuídos fisicamente;
- Uma das vantagens do modelo de *cloud* é pagar apenas pelos recursos utilizados;
- No modelo Computação em Nuvem, o risco financeiro é mensal (usa e paga) e pode-se acompanhar mais de perto como o dinheiro está sendo gasto;

- O modelo Computação em Nuvem retira da empresa todo o trabalho e o custo de administrar toda a parafernália tecnológica, que geralmente não é o seu “core business”.

Segundo Apostu *et al* (2013), as vantagens são:

- Mais fácil escalar serviços, “ela torna mais fácil para empresas escalar o serviço de acordo com as demandas dos clientes” (APOSTU *et al*, 2013, p. 120).
- Armazenamento quase ilimitado, “armazenamento de informações na nuvem é oferecido praticamente uma capacidade ilimitada” (APOSTU *et al*, 2013, p. 120).

E Reese (2009) destaca a seguinte vantagem:

- Alocação de recursos quando necessário. A possibilidade de alocar os recursos somente quando forem necessários também elimina a necessidade da aquisição antecipada de equipamentos de TI. Com isso, esse investimento pode ser utilizado em outras atividades estratégicas da empresa.

#### 2.4.2 Desvantagens

Dentre as desvantagens da Computação em Nuvem, destacamos as seguintes:

- Posse dos dados, “aplicativos e informações confidenciais do cliente estarão em posse dos provedores de nuvem” (SILVA, 2013, p. 24);
- Uma das maiores empresas americanas relatou que 53% dos profissionais estão com suas atenções redobradas para a segurança da Nuvem após a adesão das soluções ao ambiente Nuvem. Isso mostra que a segurança da informação apresenta-se como uma das desvantagens da Computação em Nuvem;
- Exposição de dados é uma grande preocupação dos profissionais de segurança, uma vez que as informações estão sujeitas a grande exposição de ataques. Toda essa preocupação demonstra uma vulnerabilidade;
- O desempenho. “Em alguns casos, é possível que determinadas aplicações fiquem lentas” (SILVA, 2013, p. 24), podendo afetar diretamente na transferência de dados e conseqüentemente no tempo de resposta dessas informações solicitadas;

- Vulnerabilidade a acesso e questões de privacidade das informações que estão na nuvem;
- Outra desvantagem que se pode citar, é o fato de que, para utilizar a Computação em Nuvem obrigatoriamente necessita-se de uma conexão constante com a internet, preferencialmente uma banda larga, pois ela não permite trabalhar sem que estejamos conectados;
- Problemas técnicos. “Embora seja verdade que informações e dados sobre a nuvem podem ser acessadas a qualquer momento e em qualquer lugar, há momentos em que o sistema pode ter alguns danos graves” (APOSTU *et al*, 2013, p. 120). Deve existir a conscientização das empresas que a tecnologia pode ser propensa a falhas e outros problemas técnicos;
- E, segundo Velte *et al.* (2010), uma outra desvantagem é a falta de interoperabilidade e portabilidade. O tópico seguinte trata sobre portabilidade da nuvem, alertando para a capacidade de adaptação dos provedores na hora de migrar seus negócios e/ou serviços.

Nesse referencial teórico também se aborda a área de conhecimento do aprendizado automático, explanando alguns conceitos de métodos de classificação e aprendizado. A seção a seguir apresenta essa discussão.

## 2.5 Aprendizado automático

Aprendizado automático ou Aprendizagem de Máquina (*Machine learning*), como também é conhecida, é uma técnica que permite que o computador aprenda ou aperfeiçoe seu desempenho em certas tarefas através de algoritmos inteligentes: trata-se de um subcampo da Inteligência Artificial.

Para Russell e Norvig (2003), a Inteligência Artificial é uma das ciências mais recentes que atualmente abrange uma variedade enorme de subcampos, que vão desde áreas de uso geral, como aprendizado e percepção, até tarefas mais específicas, como jogos de xadrez.

Geralmente utilizam-se os algoritmos de aprendizagem de máquina que identificam padrões dentro de um conjunto de dados. Alguns desses algoritmos existem há bastante tempo, sendo aplicados nas mais diferentes áreas do conhecimento, indústria, meteorologia e na medicina com diagnóstico médico.

Com o uso de aprendizado automático, faz-se necessário o estabelecimento de algumas definições em relação aos métodos de aprendizagem a serem aplicados, como tarefa de classificação por árvore de decisão com algoritmo de classificação J48, método de aprendizagem por indução. Esses conceitos são apresentados a seguir.

### 2.5.1 Classificação por Árvore de Decisão

A classificação por Árvore de Decisão é uma técnica que pode ser aplicada nos treinamentos Supervisionados<sup>8</sup> e Não-supervisionados<sup>9</sup>. Sua aplicação é com o objetivo de prever valores de variáveis do tipo categóricas que assumem valores finitos e que podem ser ordenadas. Por exemplo, pode-se criar um modelo para classificar os clientes de um estabelecimento como bom pagador ou mau pagador. Nesse cenário, a classificação estabelece as categorias existentes no treinamento. Camilo e Carlos (2009, p. 12) fazem um detalhamento desse funcionamento da classificação por árvore de decisão:

O método de classificação por Árvore de Decisão funciona como um fluxograma em forma de árvore, onde cada nó (não folha) indica um teste feito sobre um valor (por exemplo, (idade > 20)). As ligações entre os nós representam os valores possíveis do teste do nó superior, e as folhas indicam a classe (categoria) a qual o registro pertence. Após a árvore de decisão montada, para classificarmos um novo registro, basta seguir o fluxo na árvore (mediante os testes nos nós não-folhas) começando no nó raiz até chegar a uma folha. Pela estrutura que formam, as árvores de decisões podem ser convertidas em Regras de Classificação.

A árvore de decisão também pode ser gerada utilizando a configuração com podagem, utilizando o método por indução e não incremental, segundo Breiman (1998). Essa configuração é utilizada para prevenir o problema do *overfitting*, que é um ajuste demasiado dos dados de treinamento, e para melhorar a classificação deve-se realizar a poda ou não da árvore.

---

<sup>8</sup> Neste tipo de aprendizagem, existe um "professor" que avalia a resposta da rede ao padrão atual de *inputs*. Disponível em: <<http://home.isr.uc.pt/~paulo/PROJ/NN95/node31.html>>. Acesso em: 10 de dezembro de 2015.

<sup>9</sup> Nesta forma de aprendizagem, não existe "professor". A rede tem de descobrir sozinha relações, padrões, regularidades ou categorias nos dados que lhe vão sendo apresentados e codificá-los nas saídas. Disponível em: <<http://home.isr.uc.pt/~paulo/PROJ/NN95/node31.html>>. Acesso em 10 de dezembro de 2015.

Para Camilo e Carlos (2009), o sucesso das árvores de decisão se deve ao fato de ser uma técnica extremamente simples, a qual não necessita de parâmetros de configuração, tendo geralmente um grau de assertividade muito eficiente. Apesar de ser uma técnica extremamente poderosa, é necessária uma análise detalhada dos dados que serão usados para garantir bons resultados.

## **2.6 Adoção/Migração Para Computação em Nuvem**

Segundo uma pesquisa realizada pela IDC<sup>10</sup> Brasil, o mercado de Nuvem Pública somou cerca de US\$ 64 milhões (FIGUEIREDO, 2013). O crescimento médio anual projetado entre 2010 e 2014 é de 66,2%, totalizando um mercado de US\$ 491,4 milhões em 2014. Em 2010, o mercado de SaaS representou 32% (US\$ 20,8 mi) do valor total de mercado. O mercado de PaaS obteve participação de 12% (US\$ 7,7 mi) e o mercado de IaaS, 56% (US\$ 35,8 mi). Em 2014, a IDC espera que o mercado de SaaS represente 39% (US\$ 192 mi) do total, o mercado de PaaS, 8% (US\$ 39 mi) e o mercado de IaaS, 53% (US\$ 261 mi). O que todas essas informações estatísticas representam são as alternativas diversificadas para o mercado brasileiro no tocante à Computação em Nuvem.

De acordo com Verderami (2013), essas empresas buscam mais alternativas para suporte no avanço de suas atividades, sendo que um dos principais aspectos apontados como condição de desenvolvimento de suas atividades são a valorização e a inovação. Embora esses números da IDC apresentem perspectivas positivas, as barreiras para a adoção do modelo no mercado corporativo nacional estão localizadas na adaptação cultural (percentualmente, 84%), que consiste na possibilidade de as empresas habituarem-se à ideia de utilizarem os serviços em um ambiente totalmente virtualizado. O uso do modelo tradicional de infraestrutura local ainda é utilizado em larga escala dentro das empresas e essa é essa cultura que pode sofrer uma transformação com a Nuvem.

Outras variáveis dizem respeito a segurança e custo, empatados em 21%, conforme relata pesquisa realizada por Jacquet e Cavassana (2012). Entretanto, é possível que as empresas brasileiras se tornem muito mais agressivas em avançar seus planos de Nuvem após o tradicional atraso em relação aos Estados Unidos e

---

<sup>10</sup> Empresa de Consultoria em Tecnologia. Disponível em: <[www.idcbrasil.com.br](http://www.idcbrasil.com.br)>. Acesso em: 05 de janeiro de 2016.

Europa, de acordo com a pesquisa realizada pela Capgemini com 415 tomadores de decisão (MARI, 2014).

Na seção a seguir, abordam-se os desafios encontrados na migração da computação em nuvem no Brasil.

## 2.7 Desafios para essa migração no Brasil

No Brasil, podemos destacar alguns desafios que são prováveis influenciadores na migração da Nuvem, dentre eles a falta de conhecimento da tecnologia Nuvem, a segurança da informação, os provedores de banda larga e a aplicabilidade das leis e normas:

- **Conhecimento da tecnologia nuvem:** a computação em nuvem, teoricamente, trata-se de uma tecnologia recente, algumas empresas sequer se familiarizam com o termo. Uma pesquisa realizada por Jacquet e Cavassana (2012) revelou que as empresas brasileiras ainda sofrem com a falta de conhecimento sobre o tema. Diante disso, migrar para um ambiente desconhecido torna-se ainda mais difícil.

- **Segurança da Informação:** a gestão de segurança não é uma tarefa trivial, é provável que aplicar a política de segurança seja um das maiores preocupações e desafios, não só do ambiente Nuvem, mas para qualquer ambiente que possa compartilhar ou manipular informações. “A Política de Segurança da Informação – PSI é o documento de diretriz, é o primeiro passo na construção de uma gestão de segurança da informação” (LOREIRO, 2008, p. 43). Garantir que os procedimentos atrelados à política de segurança e os padrões sejam adotados é uma tarefa importante, uma vez que manter um ambiente seguro está cada vez mais difícil, principalmente para empresas que não possuem maturidade para tal. Esse desafio não cabe apenas ao ambiente Nuvem.

- **Provedores de banda larga:** outro desafio encontrado refere-se aos serviços prestados pelas operadoras de banda larga. De acordo com os dados da Anatel, o Brasil terminou janeiro de 2016 com 25,4 milhões de acessos banda larga e 12,4 acessos/100 habitantes, cerca de 85% desses acessos são realizados através das operadoras de telefonia, tendo a Claro como a maior provedora com 32,20% dessas conexões (TELECO, 2016). Existe uma deficiência desses serviços

por parte das operadoras no tocante à demanda de usuários, à velocidade de dados trafegada e à abrangência da operadora. Diante disso, como a maioria das conexões são providas pelas operadoras de telefonia, pode-se entender como um desafio a melhoria da prestação desse serviço pelas operadoras.

- **Aplicabilidade das leis e normas:** fazer com que as leis e normas sejam aplicadas não é uma atividade simples. A instrução normativa GSI/PR nº 1, de 13 de junho de 2008, disposta nos incisos de I a IX do artigo 2º, aplica-se não somente em ambientes convencionais, mas também em ambiente Nuvem. Essa instrução determina que os fornecedores de serviços de informações devem seguir as seguintes normas estabelecidas:

**I - Política de Segurança da Informação e Comunicações:** documento aprovado pela autoridade responsável pelo órgão ou entidade da administração pública federal, direta e indireta, com o objetivo de fornecer diretrizes, critérios e suporte administrativo suficientes à implementação da segurança da informação e comunicações;

**II - Segurança da Informação e Comunicações:** ações que objetivam viabilizar e assegurar a disponibilidade, a integridade, a confidencialidade e a autenticidade das informações;

**III - Disponibilidade:** propriedade de que a informação esteja acessível e utilizável sob demanda por uma pessoa física ou determinado sistema, órgão ou entidade;

**IV - Integridade:** propriedade de que a informação não foi modificada ou destruída de maneira não autorizada ou acidental;

**V - Confidencialidade:** propriedade de que a informação não esteja disponível ou revelada a pessoa física, sistema, órgão ou entidade não autorizado e credenciado;

**VI - Autenticidade:** propriedade de que a informação foi produzida, expedida, modificada ou destruída por uma determinada pessoa física ou por um determinado sistema, órgão ou entidade;

**VII - Gestão de Segurança da Informação e Comunicações:** ações e métodos que visam à integração das atividades de gestão de riscos, gestão de continuidade do negócio, tratamento de incidentes, tratamento da informação, conformidade, credenciamento, segurança cibernética, segurança física, segurança

lógica, segurança orgânica e segurança organizacional aos processos institucionais estratégicos, operacionais e táticos, não se limitando, portanto, à tecnologia da informação e comunicações;

**VIII - Quebra de segurança:** ação ou omissão, intencional ou acidental, que resulta no comprometimento da segurança da informação e das comunicações;

**IX - Tratamento da informação:** recepção, produção, reprodução, utilização, acesso, transporte, transmissão, distribuição, armazenamento, eliminação e controle da informação, inclusive as sigilosas.

Para Harauz *et al.* (2009), existe outro desafio: a confiança. Ela é a grande preocupação no modelo de Computação em Nuvem. Para armazenar dados das empresas nas nuvens, existem fatores como jurisdição, responsabilização, privacidade e ameaças associadas à tecnologia de virtualização.

Outros desafios que o governo brasileiro precisa enfrentar para a migração para a Nuvem foram revelados no levantamento do Núcleo de Pesquisa de Segurança em Computação em Nuvem:

**Perda da Governança:** o gerenciamento dos riscos em segurança da informação gera complexidade, tanto para quem contrata como para quem fornece, na segurança de dados e informações, nas questões legais e contratuais, na auditoria e conformidade e na interoperabilidade e portabilidade.

**Segurança técnica e operacional da Nuvem:** implica na segurança de acessos físicos e lógicos, na virtualização, nas respostas a incidentes e nas análises forense, pois possui complexidade para execução de perícias e obtenção de evidências, desafios no gerenciamento de chaves criptográficas e, por fim, na segurança de aplicações com gerenciamento de mudanças e detecção de códigos maliciosos.

**Desafios da legislação brasileira:** versam sobre contratação de serviços de Tecnologia da Informação: SLTI/MP (IN-04), bem como o acesso à informação regido pela lei de nº 12.527 e decreto nº 7.724. A lei 12.527 dita em seu artigo 3º:

Os procedimentos previstos nesta Lei destinam-se a assegurar o direito fundamental de acesso à informação e devem ser executados em conformidade com os princípios básicos da administração pública e com as seguintes diretrizes:

- I - observância da publicidade como preceito geral e do sigilo como exceção;
- II - divulgação de informações de interesse público, independentemente de solicitações;



- III - utilização de meios de comunicação viabilizados pela tecnologia da informação;
- IV - fomento ao desenvolvimento da cultura de transparência na administração pública;
- V - desenvolvimento do controle social da administração pública.

Vale ressaltar que à lei 12.527 subordinam-se os órgãos públicos, autarquias, fundações públicas, empresas públicas, sociedades de economia mista, entidades controladas pela União, assim como entidades sem fins lucrativos que recebem recursos públicos.

## **2.8 Resumo do capítulo**

Neste capítulo, abordou-se o referencial teórico dessa dissertação. Foram contextualizados os principais conceitos e definições da Computação em Nuvem e as tecnologias habilitadoras da Computação em Nuvem, assim como as características essenciais da Nuvem de acordo com o NIST, e, por fim, alguns conceitos sobre aprendizagem automática e classificação por árvore de decisão. Apresentou-se também o processo de Adoção dessa arquitetura e finalizando com a abordagem dos principais desafios enfrentados pelas empresas brasileiras na adoção e permanência na Nuvem. O próximo capítulo aborda a metodologia utilizada para o desenvolvimento da pesquisa e detalhes relacionados aos meios de divulgação, público alvo, pré-teste, ferramentas utilizadas e planejamento amostral.

### 3 METODOLOGIA DA PESQUISA

Este capítulo explana a metodologia da pesquisa, aplicando os conceitos e métodos utilizados neste estudo.

Na seção 3.1 apresentam-se as definições metodológicas e suas especificidades aplicadas na pesquisa; a seção 3.2 cita as definições realizadas no aprendizado automático que foram aplicados nos treinamentos; na seção 3.3, apresenta-se o propósito da pesquisa; na seção 3.4, encontra-se planejamento amostral da pesquisa sob três perspectivas: dados preliminares, empresas que não aderiram à Nuvem e empresas que já estão utilizando algum serviço em Nuvem; na seção 3.5, são apresentados público alvo, pré-teste e meios de divulgação desta pesquisa; a seção 3.6 aborda os recursos utilizados, como ferramenta de implementação de pesquisa, Servidor EC2 da AWS, Gerenciador de Serviços, SGDB, ferramenta de estatísticas; por fim, a seção 3.7 apresenta o resumo do capítulo.

A pesquisa realizada baseia-se em um estudo empírico e criterioso sobre Computação em Nuvem a fim de apresentar um quadro representativo da realidade brasileira em relação à migração e à satisfação acerca dessa arquitetura. Mais especificamente, a pesquisa constitui-se como exploratória de natureza descritiva e explicativa, buscando dados e informações sobre sistema de Computação em Nuvem. Para identificação desses aspectos, desenvolveu-se uma pesquisa (*Survey*), que se efetiva buscando na realidade dados e informações sobre alguns questionamentos no contexto da Computação em Nuvem. Inicia-se esse capítulo com a seção 3.1, Definição metodológica, que trata dos métodos aplicados no desenvolvimento da pesquisa.

#### 3.1 Definição metodológica

Para o desenvolvimento dessa pesquisa dissertativa foram definidas algumas estratégias metodológicas com o propósito de aplicar o melhor método para esta investigação. Discorre-se a seguir acerca do tipo de pesquisa, tipo de abordagem, procedimento técnico, instrumento aplicado, forma de coleta e tipo de análise dos resultados:

**Tipo de pesquisa:** exploratória de natureza descritiva e explicativa, tendo “como finalidade proporcionar mais informações sobre o assunto que vamos investigar, possibilitando sua definição e seu delineamento” (CRISTIANO; CESAR, 2013, p. 51). “A pesquisa de natureza descritiva é, juntamente com as pesquisas exploratórias, as que habitualmente realizam os pesquisadores sociais preocupados com a atuação prática” (CRISTIANO; CESAR, 2013, p. 53). A pesquisa explicativa é aplicada quando o pesquisador procura explicar os porquês dos fenômenos e suas causas, por meio do registro, da análise, da classificação e da interpretação dos fenômenos observados.

**Ênfase da abordagem:** abordagem quantitativa, uma vez que faz uso de dados estatísticos para a interpretação dos resultados coletados. “Considera o que pode ser quantificável, o que significa traduzir em números opiniões e informações para classificá-las e analisá-las” (KAUARK *et al.*, 2010, p. 27).

**Procedimento técnico:** levantamento através de um *Survey*, é considerado um dos elementos mais importantes para a identificação de um delineamento, é um procedimento adotado para a coleta de dados (CRISTIANO; CESAR, 2013). Existem dois grupos de delineamentos: fontes de papel e aqueles cujos dados são fornecidos por pessoas, sendo que este último foi adotado como procedimento para a realização dessa pesquisa do tipo “pesquisa com participantes”.

**Instrumento aplicado:** questionário com 14 (quatorze) itens. “O questionário é uma série ordenada de perguntas que devem ser respondidas por escrito pelo informante (respondente)” (CRISTIANO; CESAR, 2013, p. 108). É considerado um instrumento ou programa de coleta de dados que utiliza uma linguagem simples e direta para facilitar o entendimento dos participantes.

**Forma de coleta de dados:** feita através de um formulário *WEB (Online)*, aplicando-se o conceito de interrogação direta dos respondentes por meio do preenchimento de formulário, visando uma maior abrangência proporcionada pela internet. A coleta de dados consiste na “seleção e organização do material coletado – tabulação de dados, análise e discussão dos resultados, conclusão da análise” (KAUARK *et al.*, 2010, p. 29). Sabe-se da existência de outro instrumento de coleta como aplicação da entrevista face a face, porém esse método não foi utilizado para esta pesquisa porque poderia interferir na quantidade de participantes pretendidas na pesquisa. Entende-se por entrevista “a obtenção da informação do entrevistado sobre determinado assunto ou problema”. (CRISTIANO; CESAR, 2013, p. 108).

**Tipo de Análise dos resultados:** para a análise dos resultados a partir dos dados coletados utiliza-se o aprendizado automático, também conhecido como aprendizagem de máquina, a fim de treinar a base de dados gerada com o propósito de encontrar a melhor classificação das informações propostas nos objetivos da pesquisa.

A seção 3.2 a seguir detalha o funcionamento desse processo metodológico.

### 3.2 Métodos do aprendizado automático

Para aplicar aprendizado automático é necessário que sejam estabelecidas algumas definições em relação aos métodos a serem utilizados para possibilitar que o treinamento apreenda, por exemplo, a tarefa a ser realizada, o método de aprendizagem, o modo de treinamento, a hierarquia do aprendizado, o paradigma do aprendizado, o tipo de treinamento e as variáveis de classificação (atributo classificador). Nesse sentido, são essas as definições importantes para esse trabalho:

**Tarefa de treinamento:** classificação por árvore de decisão. Essa tarefa pode ser aplicada para prever os rótulos e as classes do conjunto de exemplos e classificá-las de acordo com o objetivo do aprendizado. “As árvores de decisão constituem uma técnica muito poderosa e amplamente utilizada em problemas de classificação” (ZUBEN; ATTUX, 2010, p. 10).

**Método de aprendizagem:** método por indução. Esse aprendizado geralmente se apresenta como supervisionado e não supervisionado (CHAPELLE *et al.*, 2006). Pode-se dizer que é uma inferência de uma regra geral (hipóteses) a partir de exemplos particulares que induzem a uma conclusão.

**Modo de Treinamento:** modo não incremental. Nesse modo de treinamento “é necessário que todos os exemplos de treinamento estejam disponíveis simultaneamente para a indução do modelo de aprendizagem do classificador” (STANGE; NETO, 2011, p. 36), pois não permite adicionar um novo exemplo a partir do conjunto de treinamento existente. O contrário disso é o modo incremental, que permite a cada nova interação o acréscimo de um novo exemplo ao conjunto de treinamento processado.

**Hierarquia do aprendizado:** hierarquia de aprendizado Supervisionado. “Ele assume a existência de um professor, alguma medida de adequação ou outro

método externo de classificação de exemplos de treinamento” (LUGER, 2007, p. 372). A hierarquia de aprendizado supervisionada é bastante utilizada nas tarefas de classificações.

**Paradigma do aprendizado:** paradigma simbólico. "Os algoritmos de árvores de decisão proporcionam uma das melhores abordagens metodológicas à aquisição de conhecimento simbólico" (DE PAULA, 2002, p. 10), pois consiste em realizar treinamento com mudanças adaptativas do sistema no sentido de lhe permitir realizar a tarefa de forma mais eficiente e efetiva.

**Tipo de treinamento:** configuração com poda. A realização do treinamento ajustado com poda nada mais é do que ajustes aplicados aos treinamentos que podem gerar melhor resultados e conseqüentemente os melhores casos classificados, visto que “a árvore podada se torna mais simples, facilitando a sua interpretabilidade por parte do usuário” (ZUBEN; ATTUX, 2010, p. 31), eliminando as subárvores consideradas insignificantes para o resultado.

**Variáveis classificadoras:** SIM/NÃO e SATISFEITO/INSATISFEITO. “A tarefa do classificador é realizar um mapeamento dos atributos para um diagnóstico” (ZUBEN; ATTUX, 2010, p. 7). Para estabelecer o atributo classificador nos treinamentos, foram atribuídos “SIM”, com o objetivo de representar participantes propensos a contratar os serviços em Nuvem, e “NÃO”, para aqueles propensos a não contratar os serviços em nuvem. Para aqueles respondentes que já se encontram utilizando Computação em Nuvem, definiram-se os atributos “SATISFEITO” e “INSATISFEITO” como identificação do nível de satisfação por estar na Nuvem.

Para essa classificação utilizou-se neste trabalho o algoritmo J48, embora existam outras técnicas/tarefas com outros algoritmos, como *cluster* com algoritmo CobWeb, associação com algoritmo Apriori, entre outros. A justificativa da utilização desse algoritmo é que, além de ser um classificador comumente utilizado pela comunidade acadêmica, é possível trabalhar com variáveis categóricas, além de viabilizar fácil interpretação dos resultados. Ele também é um classificador baseado em árvore de decisão que pode mostrar/evidenciar as relações que levam um determinado respondente a utilizar ou não Nuvem. Outro fator motivador para usar a tarefa de classificação é a possibilidade de criar um modelo capaz de, a partir dos questionamentos levantados, as perguntas comuns, ser possível prever as

evidências de utilização ou não da Computação em Nuvem, bem como o seu nível de satisfação.

### **3.3 Propósitos da Pesquisa**

A proposta desta pesquisa é a realização de um estudo planejado por meio de uma pesquisa exploratória de natureza descritiva e explicativa com abordagem quantitativa. A finalidade é encontrar indícios pelos quais algumas empresas podem estar mais propensas a migrar seus serviços e/ou infraestruturas de TIC para o sistema de Computação em Nuvem no Brasil. Adicionalmente, para aquelas que já adotaram esta arquitetura em nuvem, investigar se já têm condições de se afirmarem satisfeitas ou não e por quais motivos. O propósito é encontrar um padrão de resposta por meio do aprendizado automático, com base no modelo tradicional (infraestrutura local). Sua aplicação acontece no cenário das empresas de pequeno, médio e grande porte, de acordo com a classificação do BNDES (2010).

Os questionamentos são baseados em 3 (três) perspectivas de investigação: porte e atuação das empresas; características específicas da Nuvem; e segurança.

Para a investigação do porte e atuação das empresas, o propósito é averiguar a dimensão da empresa, tendo em vista traços como faturamento, segmento e número total de colaboradores, para possibilitar uma visão aproximada da classificação dessas empresas (pequeno, médio ou grande porte). Os questionamentos quanto às características específicas da Nuvem versam sobre modelo de serviço, tipos de nuvem e redução de custo. A questão da Segurança da Nuvem pública é tratada com o intuito de identificar se a segurança associada a outras características são fatores determinantes nessa decisão, tanto satisfação ou insatisfação, contratação ou não, na perspectiva das empresas que usam ou não a Nuvem. Baseadas nessas linhas de investigação, foram elaboradas 14 (quatorze) questões, conforme a seção 3.4 apresenta a seguir.

### **3.4 Planejamento Amostral (Questionamentos)**

Esta pesquisa foi definida a partir de 14 (quatorze) questionamentos, em 3 (três) perspectivas: i) 5 (cinco) questionamentos referentes aos dados preliminares

da empresa ou organização; ii) 4 (quatro) questionamentos para as organizações que não utilizam serviços em nuvem; iii) e 5 (cinco) questionamentos para as organizações que já utilizam algum tipo de serviços em nuvem. Os questionamentos de cada perspectiva são apresentados nas subseções a seguir.

#### 3.4.1 Perspectivas da identificação da Empresa ou Organização (I):

1. Qual a área de atuação da sua Empresa/Organização?

Alternativas: Comércio Varejista; Comércio Atacadista; Prestação de Serviços; Finanças; Industrial; Pública-Privada; Outros.

2. Qual o faturamento atual bruto (anual) da sua Empresa/Organização?

Alternativas: Menos que 500 Mil; Entre 500 mil e 1 milhão; Entre 1 milhão e 2 milhões; Entre 2 milhões e 10 milhões; Entre 10 milhões e 1 bilhão; Acima de 1 bilhão.

3. Qual o número de funcionários de sua Empresa/Organização?

Alternativas: Menor que 20 colaboradores; Entre 21 e 50 colaboradores; Entre 51 e 100 colaboradores; Entre 101 e 200 colaboradores; Entre 201 e 500 colaboradores; Entre 501 e 1000 colaboradores; Acima de 1000 colaboradores.

4. Você é o responsável pela TIC na empresa?

Alternativas: Sim/Não

5. Sua Empresa/Organização utiliza algum tipo de serviço da tecnologia em nuvem?

Alternativas: Sim/Não

#### 3.4.2 Perspectivas das empresas que ainda não aderiram à nuvem (II):

1. Qual tipo de Nuvem sua Empresa/Organização contrataria?

Alternativas: Nuvem Privada; Nuvem Pública; Nuvem Comunitária; Nuvem Híbrida.

2. Das características abaixo, indique qual a sua Empresa/Organização optou pela Nuvem?

Alternativas: Confiabilidade; Segurança da informação; do serviço; Inoperabilidade; Manutenibilidade; Garantia de entrega; Outros.

3. Se sua Empresa/Organização resolvesse migrar para a nuvem, qual serviço contrataria?

Alternativas: Plataforma PaaS; Infraestrutura IaaS; Software SaaS; Armazenamento STaaS; Forensic FaaS; Banco de dados como Serviço DBaaS; Outros.

4. Qual empresa e serviço sua Empresa/Organização contrataria?

Alternativas: Amazon; Azure; Google; HP; Intel; Sun; 3Tera; Outras.

### 3.4.3 Perspectivas das empresas que fazem uso da tecnologia nuvem (III):

1. Qual serviço sua Empresa/Organização encontra-se utilizando?

Alternativas: Software como Serviço SaaS; Infraestrutura como Serviço IaaS; Plataforma como Serviço PaaS; Armazenamento como Serviço STaaS; Forensic como Serviço FaaS; Banco de dados como Serviço DaaS; Outros.

2. Das características abaixo, indique qual a sua Empresa/Organização optou pela Nuvem?

Alternativas: Disponibilidade; Escalabilidade; Segurança; Qualidade do serviço; Preços; Outros motivos.

3. Que tipo de Nuvem sua Empresa/Organização utiliza?

Alternativas: Nuvem Privada; Nuvem Pública; Nuvem Comunitária; Nuvem Híbrida.

4. Em caso de Nuvem Pública, sua Empresa/Organização sente-se confortável quanto à segurança na sua utilização?

Alternativas: Pouco confortável; Muito confortável; Razoavelmente confortável; Ligeiramente desconfortável; Totalmente desconfortável; Nenhuma alternativa.

5. Qual o percentual de redução de custo que sua empresa alcançou com a migração para a nuvem?

Alternativas: Menos que 20%; Aproximadamente 20%; Aproximadamente 30%; Aproximadamente 40%; Acima de 50%; Não houve redução de custo.

O planejamento amostral foi definido com base nas linhas de investigação abordadas na seção 3.3. Após essa definição, cabe delimitar o público-alvo, os pré-



testes que foram realizados, assim como os meios de divulgação para proporcionar a coleta de dados da pesquisa.

### 3.5 Público-alvo, Pré-testes e Meios de divulgação

A pesquisa foi direcionada para alguns segmentos, sendo encaminhado convite de participação da pesquisa para TIC (Associações, Indústria, Comércio e Tecnologia), Instituições de Ensino Superior – IES (públicas e privadas) e os Comércios (Varejistas e atacadistas). Utilizou-se o espaço amostral de tamanho 20 (vinte) como pré-teste, entre os dias 1º e 10 de fevereiro de 2015, antes da liberação definitiva para os respondentes. Assim, identificou-se que seriam necessários alguns ajustes em algumas questões para melhor entendimento dos usuários respondentes. Essas correções foram realizadas com o *feedback* enviado ao e-mail dos respondentes, contendo os seguintes questionamentos:

- Quanto tempo você levou para responder o questionário?
- As questões estão dispostas em uma ordem lógica?
- Existe algum termo que não esteja apresentado com clareza?
- Os conceitos abordados estão bem detalhados?
- Se você fosse modificar alguma questão ou alternativa, qual modificaria?

Após esse mapeamento, identificou-se a necessidade de algumas alterações, como por exemplo, utilizar o uso do condicional: a questão só é exibida mediante a resposta de uma questão anterior. Também foi sugerida uma melhor ordenação nas perguntas. Algumas questões alinhavam-se melhor em uma determinada posição ou ordem. A ordem das questões também otimizou o tempo utilizado para responder ao questionário, que requeria em média de 4,5 minutos.

Após a realização dos testes, buscou-se um meio de divulgação que possibilitasse uma boa abrangência, para isso foram utilizados recursos da *web* (e-mail, blog, redes sociais, página web) com o intuito de divulgar esta pesquisa em diversas regiões do país, em especial por meio de redes sociais, tendo o *LinkedIn* como ferramenta mais utilizada. Uma pesquisa realizada em 2014 pela agência

Americana de mídias sociais *WE ARE SOCIAL*<sup>11</sup> divulgou um status da Internet pelo mundo que reuniu e cruzou dados de diversas pesquisas de diferentes órgãos. Identificou-se que 49% da população brasileira encontra-se conectada de alguma forma à Internet, dentre os quais 43% encontram-se nas redes sociais. Daí o motivo pelo qual se escolheu as redes sociais como instrumento de divulgação da pesquisa.

### 3.6 Recursos utilizados (Ferramentas)

No desenvolvimento de uma pesquisa dessa natureza (*online*), podem se utilizar diversos recursos na implementação da pesquisa. Para garantir que ela tivesse disponibilidade de acesso via Internet, utilizou-se recursos do tipo: Servidor EC2 da AWS, Gerenciador de serviços XAMPP, IDE de implementação *Limesurvey*, Sistema de Gerenciamento de Banco de Dados *MySQL Server*, Ferramenta de estatística da IBM SPSS, Ferramenta de *WEKA* e Ferramenta de modelagem *dia.exe*. Descrevemos cada um desses recursos a seguir.

**Servidor EC2 da AWS:** a *Amazon.com* é uma multinacional americana de comércio eletrônico e provedora de serviço de internet, fundada nos Estados Unidos em 1994 por *Jeffrey Bezos*, ao longo do tempo consolidando-se como uma das maiores empresas no ramo de provisão de serviços em Nuvem. A plataforma EC2 é um servidor em nuvem da *Amazon*, gratuito por um ano, e foi utilizado para disponibilizar a pesquisa (Link: <<https://aws.amazon.com/pt/about-aws/>>. Acesso em 05 de fevereiro de 2016).

**Gerenciador de serviços XAMPP:** para garantir que a página da pesquisa que foi desenvolvida estivesse acessível através de outros hosts<sup>12</sup>, fez-se necessário a configuração de um servidor web (XAMPP) para possibilitar esse acesso e compartilhamento da pesquisa. O XAMPP é um gerenciador de multisserviço e de código aberto que possui os serviços gerenciamento como: Banco de dados (SGDB) MYSQL, Protocolo de Transferência de arquivo-FTP, Servidores Web, Suporte à Linguagem PHP e Perl, FileZilla e Mercury, entre outros serviços do Sistema Operacional (Link: <[www.xampp.com](http://www.xampp.com)>. Acesso em 06 de fevereiro de 2016).

---

<sup>11</sup> Agencia Americana que promove pesquisas em mídias sociais. Disponível em: <<http://wearesocial.com.br/>>. Acesso em: 15 de outubro de 2015.

<sup>12</sup> Máquina conectada a uma rede de computadores.

**IDE de desenvolvimento LIMESURVEY:** o *LimeSurvey* é uma IDE de desenvolvimento de pesquisas online, desenvolvida na linguagem de programação PHP, pelo desenvolvedor australiano *Jason Cleeland*, em 2007, publicada com o nome de *LimeSurvey* (Link: <<https://www.limesurvey.org/about-limesurvey/team-contributors>>. Acesso em: 2 de fevereiro de 2016). A utilização dessa ferramenta foi necessária para desenvolver a pesquisa propriamente dita, pois, além dos recursos de implementação, ela é prática e de fácil utilização.

**Sistema de Gerenciamento de Banco de Dados MySql Server:** a ferramenta de armazenamento de dados utilizada foi o MYSQL Server, um sistema de gerenciamento de banco de dados gratuito e de código aberto, desenvolvido na Suécia pela equipe de desenvolvedores *David Axmark, Allan Larsson e Michael Widenius*, na década de 1980, recentemente vendido à *Oracle Corporation* (Link: <<https://www.mysql.com/why-mysql/>>. Acesso em: 6 de fevereiro de 2016). Sua utilização tem como objetivo armazenar os dados e manipulá-los de forma prática e eficiente.

**Ferramenta de estatística da IBM SPSS:** o Pacote Estatístico para Ciências Sociais (SPSS) é um *software* com propósitos de análises estatísticas de dados. Possui um ambiente de fácil utilização e agradável, foi desenvolvido na Universidade de Chicago nos Estados Unidos pela IBM pela equipe de *Norman H. Nie*, tendo suas primeiras versões publicadas em 1968 (PEREIRA; PATRÍCIO, 2013). A ferramenta foi utilizada para a geração das frequências através dos gráficos gerados na fase de coleta dos dados.

**Ferramenta de WEKA:** é uma ferramenta de aprendizagem automática. O WEKA – *Waikato Environment for Knowledge Analysis* – é um *software* livre do tipo *Open Source* (código aberto) para o aprendizado automático, implementada em Java e que foi desenvolvida por um grupo de pesquisadores da Universidade de *Waikato* na Nova Zelândia. (Link: <<http://www.ibm.com/developerworks/br/opensource/library/os-weka1/>>. Acesso em 8 de fevereiro de 2016). O uso da ferramenta WEKA nessa pesquisa tem como intuito a análise e o aprendizado automático dos dados coletados, e posteriormente classificação das hipóteses.

**Ferramenta de modelagem DIA:** é um *software* livre voltado para modelagem e diagramação. Possui objetos especiais os quais ajudam a desenvolver artefatos como diagrama de classe, diagrama de entidade relacionamento, diagrama

de componente, diagrama de sistema, diagrama de fluxograma de dado, diagrama de rede e outros (Link: <<http://live.gnome.org/Dia>>. Acesso em: 03 de fevereiro de 2016). Após os treinamentos realizados pela ferramenta WEKA, é gerada uma árvore de decisão cujo formato não possui dimensões ideais, daí a utilização da ferramenta DIA para a reconstrução dessas árvores dimensionando adequadamente.

### **3.7 Resumo do Capítulo**

Neste capítulo foi apresentada a metodologia científica utilizada para o desenvolvimento deste trabalho. Para essa pesquisa, adotou-se um estudo planejado, definições metodológicas e definições do aprendizado automático, descrevendo tipo de pesquisa, ênfase da abordagem, procedimento técnico, instrumento utilizado, forma de coleta e tipo de análise aplicada. Também foram citados neste capítulo: linhas de investigação, planejamento amostral, público-alvo, pré-teste e meios de divulgação desta pesquisa. Por fim, abordaram-se as principais ferramentas utilizadas que viabilizaram a aplicação da pesquisa. O capítulo seguinte aborda a coleta dos dados, apresentando gráficos e frequências geradas na pesquisa. Essas informações possibilitaram o aprendizado automático através das análises dos resultados.

## 4 COLETA DOS DADOS

Este capítulo tem como objetivo apresentar as frequências e gráficos coletados a partir das respostas dos participantes.

Na seção 4.1 são apresentados os dados preliminares com informações que possibilitam a identificação do segmento, como faturamento e total de colaboradores das empresas ou organizações; a seção 4.2 apresenta as frequências e gráficos sob a perspectiva das empresas ou organizações que não utilizam Computação em Nuvem; na seção 4.3, citam-se as frequências e gráficos sob a perspectiva das empresas ou organizações que utilizam Computação em Nuvem; e, por fim, a seção 4.4 apresenta o resumo do capítulo.

Para viabilizar as respostas, o questionário ficou disponível aos respondentes durante os meses de março e abril de 2015. O objetivo foi atingir tanto as empresas que não utilizam a Computação em Nuvem quanto as que já utilizam algum tipo de serviço em Nuvem.

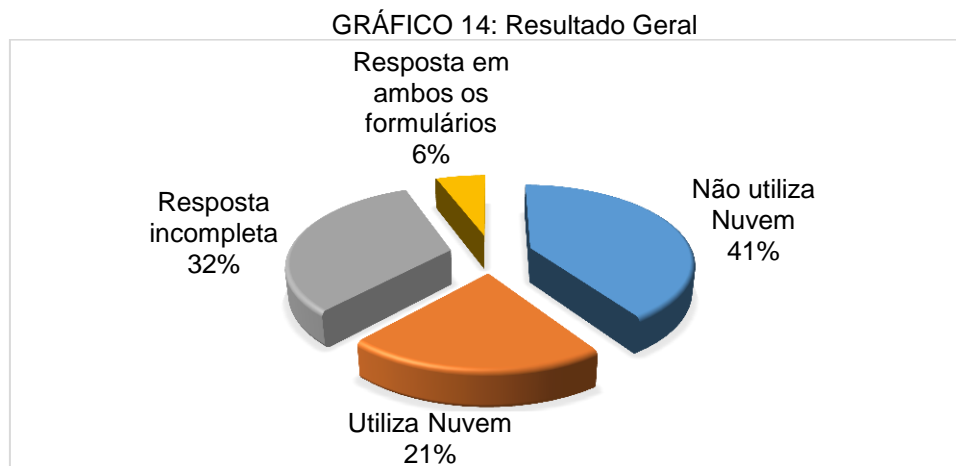
Foram coletadas informações de 432 (quatrocentos e trinta e dois) respondentes durante o período em que o estudo ficou disponível: 175 (cento e setenta e cinco), 41% dos respondentes, responderam o formulário correspondente aos que não utilizam Nuvem, 92 (noventa e dois), 21% dos respondentes, afirmaram utilizar algum tipo de tecnologia em Nuvem; 25 (vinte e cinco), 6%, responderam ambos os formulários; e 140 (cento e quarenta), 32%, responderam de forma incompleta. Não será aplicada a análise do material coletado desses dois últimos grupos, partindo do princípio que não é possível identificar a resposta correta.

Por meio dessa coleta, pode-se ter uma visão aproximada dos indícios que distanciam essas instituições da Nuvem, assim como a identificação de algumas variáveis de satisfação entre as instituições que já utilizam a Computação em Nuvem. Os resultados foram distribuídos de acordo com as perspectivas definidas na metodologia da pesquisa.

O tipo de questionário utilizado é o Estruturado e Fechado, na ordem de preenchimento dos seguintes formulários:

- DADOS PRELIMINARES;
- EMPRESAS QUE NÃO UTILIZAM COMPUTAÇÃO EM NUVEM;
- EMPRESAS QUE UTILIZAM COMPUTAÇÃO EM NUVEM.

O Gráfico 1 permite uma visão geral dessa coleta. Para possibilitar a coleta dos dados, utilizou-se a ferramenta SPSS<sup>13</sup> (Ferramenta de Estatística da IBM) para geração das estatísticas e frequências referente aos dados brutos coletados.



Fonte: próprio autor.

#### 4.1 Dados preliminares

A apresentação dos dados preliminares está subdividida quanto ao teor do questionamento efetuado, conforme segue apresentado.

##### 4.1.1 Segmento de atuação.

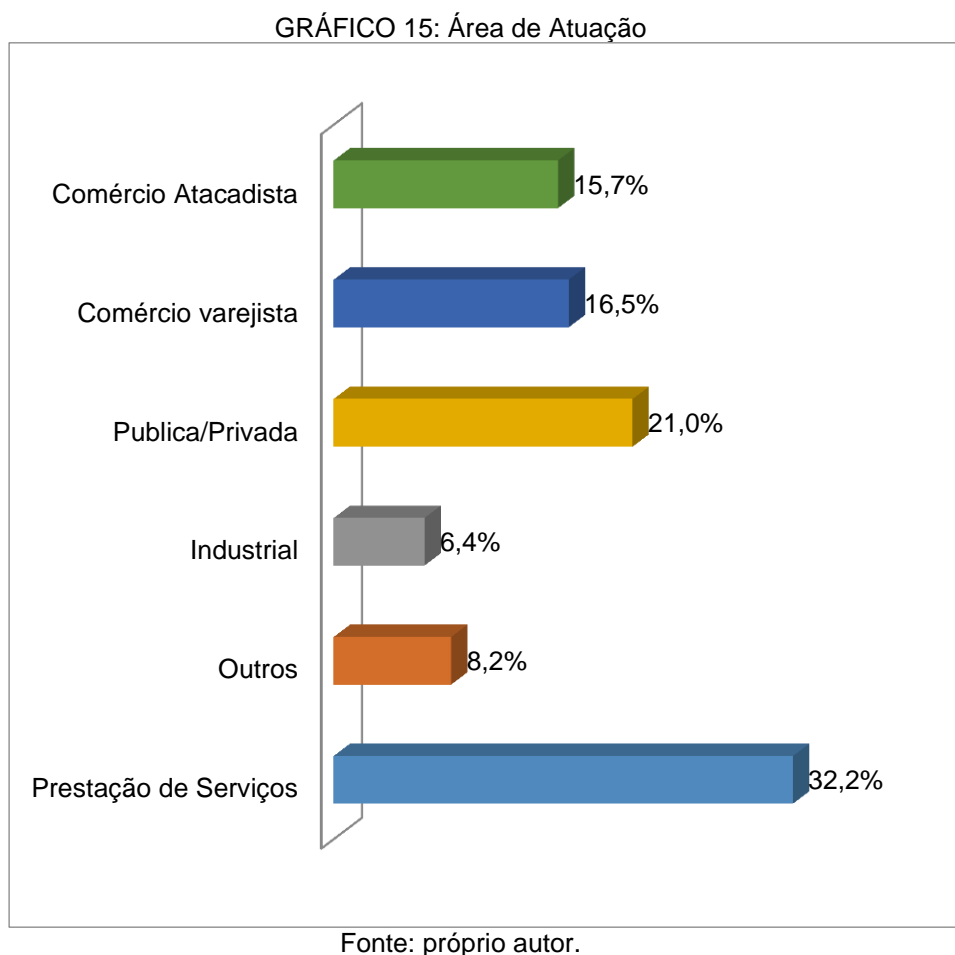
Fazer um primeiro questionamento sobre o segmento de atuação permite identificar quais os segmentos existentes entre os participantes da pesquisa. Assim, a primeira questão feita foi: *qual a área de atuação das Empresa/Organização?*

Nesse item, os seguintes resultados foram coletados: as empresas que atuam no segmento de “Prestação de Serviço” (terceirização) obtiveram o maior número de respostas, com um total de 86 respostas (32,20%); em seguida, o segmento “Pública-Privada”, com 55 respondentes (21,0%); o segmento de “Comercio Varejista” aparece em terceiro, com 44 respondentes (16,50%); em seguida, o “Comércio Atacadista” com 42 respondentes (15,70%); os “Outros segmentos”

<sup>13</sup> Statistical Package for the Social Sciences - pacote estatístico para as ciências sociais. Disponível em: <[www.ibm.com/software/products/pt/spss-stats-base](http://www.ibm.com/software/products/pt/spss-stats-base)> Acesso em: 01 de outubro de 2015.

tiveram um total de 22 respondentes (8,20%); a “Indústria” teve 17 respondentes (6,40%); e, finalmente, a área de “Finanças” não apresentou respondente.

O Gráfico 2 a seguir permite uma visão mais detalhada do resultado dessa coleta.



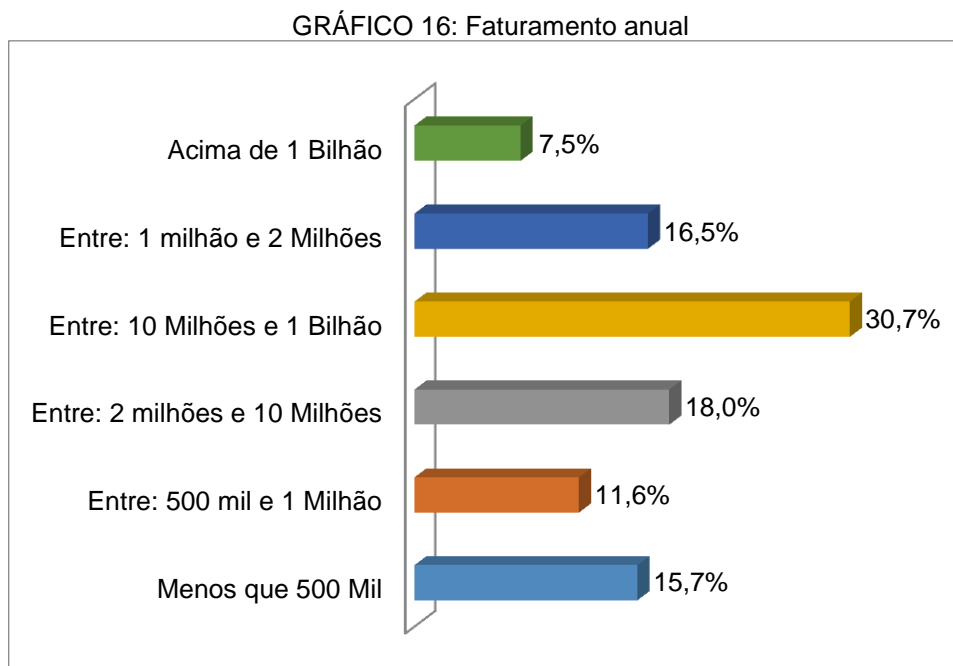
#### 4.1.2 Quanto ao faturamento anual

Identificou-se qual o faturamento anual estabelecendo-se um intervalo de valores para as empresas respondentes. A questão feita foi: *hoje, qual o faturamento bruto (anual) da sua Empresa/Organização?*

Com 82 respondentes (30,70%), encontram-se as empresas com o faturamento “Entre 10 milhões e 1 bilhão”; 48 respondentes (18,00%) declararam que suas empresas faturam “Entre 2 e 10 milhões”; 44 respondentes (16,50%) declararam que seu faturamento está “Entre 1 e 2 milhões”; de acordo com 42 respondentes (15,70%), o faturamento é “Menor que 500 mil”; 31 respondentes

(11,60%) declararam que seu faturamento se encontra “Acima de 500 mil e 1 milhão”; com 20 respondentes (7,50%) estão as empresas que faturam “Acima de 1 bilhão” de reais.

Temos uma representação mais detalhada pode ser vista no Gráfico 3.



Fonte: próprio autor.

#### 4.1.3 Quadro efetivo (Quantidade de colaboradores)

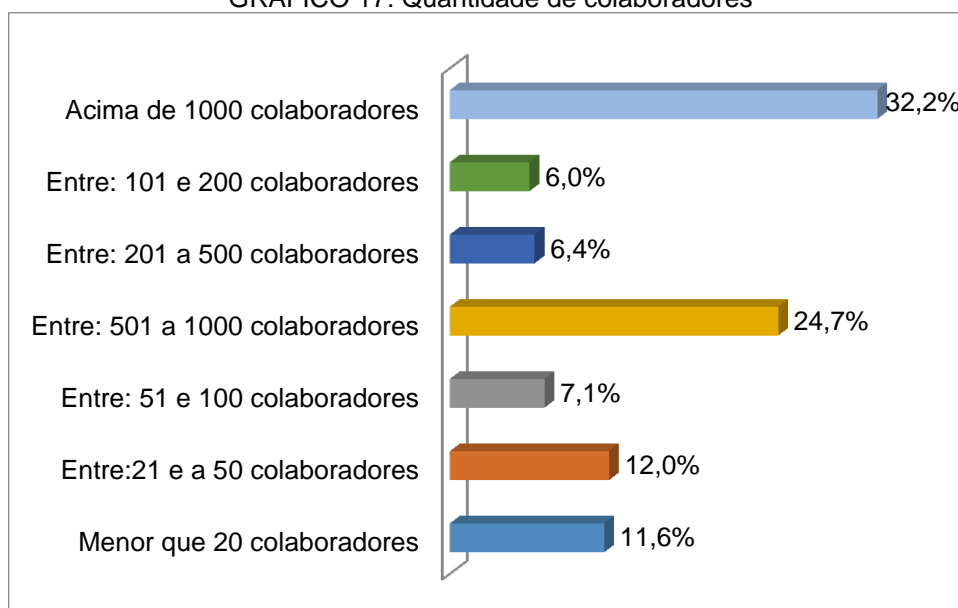
Em linhas gerais, a quantidade de colaboradores tem o propósito de identificar a quantidade de colaboradores que as empresas ou organizações possuem. Nesse sentido, a questão que se fez foi: *qual o número de funcionários de sua Empresa/Organização?*

Coletou-se que 86 respondentes (32,20%) possuem “Acima de 1000 colaboradores”; 66 respondentes (24,70%) encontram-se “Entre 501 a 1000 colaboradores”; 31 respondentes (11,60%) possuem “Menos que 20 colaboradores”; 32 respondentes (12,00%) estão “Entre 21 e 50 colaboradores”; 19 respondentes (7,10%) possuem “Entre 51 e 100 colaboradores”; 17 respondentes (6,40%) estão “Entre 201 e 500 colaboradores”; e 16 respondentes (6,00%) encontram-se “Entre 101 e 200 colaboradores”.

O Gráfico 4 apresenta os resultados desse questionamento.



GRÁFICO 17: Quantidade de colaboradores



Fonte: próprio autor.

#### 4.1.4 Responsável pela TIC

Para este item do questionário, foi solicitada a identificação do participante com o propósito de garantir que a pesquisa fosse respondida por um gestor de TIC. A questão feita foi: *você é o responsável pela TIC na empresa?*

190 respondentes (71,16%) afirmaram que “não”, enquanto que 77 respondentes (28,84%) disseram que “sim”, que são os responsáveis pelas decisões relacionadas a investimentos na área de TIC.

#### 4.1.5 Quanto à utilização da Nuvem

A partir dessa abordagem, é possível separar um total de 267 respostas que serão utilizadas para a análise da pesquisa sobre aqueles que utilizam Computação em Nuvem e aqueles que não utilizam Computação em Nuvem. A questão feita foi: *sua Empresa/Organização utiliza algum tipo de serviço da tecnologia em Nuvem?*

175 dos respondentes (65,54%) afirmaram que não utilizam Computação em Nuvem, enquanto que 92 respondentes (34,46%) afirmaram que utilizam algum tipo de serviço da Computação em Nuvem.

## 4.2 Sob a perspectiva das Empresas ou Organizações que NÃO utilizam Computação em Nuvem

Para direcionar as respostas dos respondentes que não utilizam a Computação em Nuvem, desenvolveu-se um formulário específico para esse caso, os quais podem ser visualizados a partir do tópico 4.2.1 a seguir.

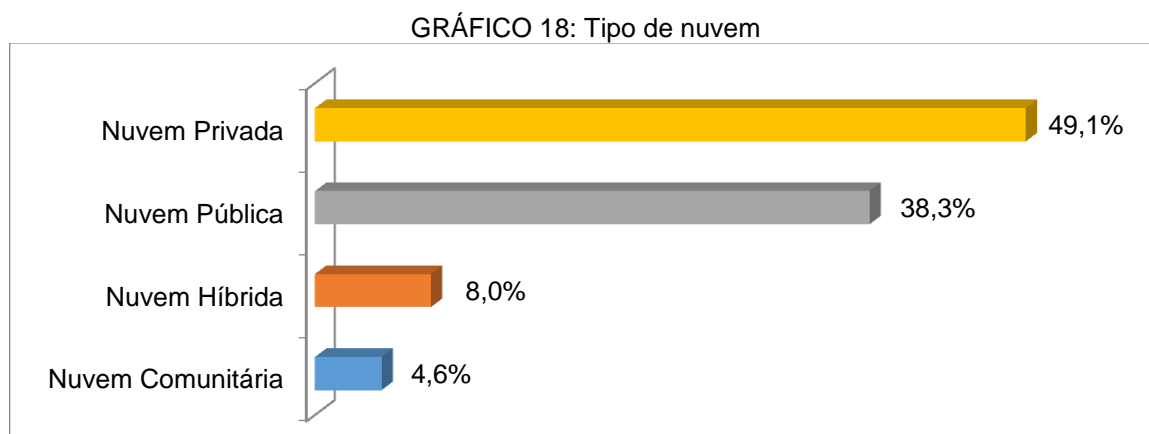
### 4.2.1 Quanto ao tipo de nuvem

O tipo de nuvem pode ser visto como um serviço específico para determinada funcionalidade que o contratante pode adquirir. Aqueles contratantes cujo objetivo é garantir a confidencialidade da informação podem optar por Nuvem Privada. Já aqueles com o propósito de cooperação podem optar por Nuvem Comunitária. Aos que têm características voltadas ao comércio, a opção poderia ser uma Nuvem Pública. E, caso o contratante queira um ambiente diversificado, pode optar por Nuvem Híbrida.

Nesse contexto, a questão feita foi: *qual tipo de Nuvem sua Empresa/Organização adquiriria?*

86 dos respondentes (49,10%) responderam que adquiririam “Nuvem Privada”; 67 respondentes (38,30%) optaram por “Nuvem Pública”; “Nuvem Híbrida” obteve 14 respostas (8,00%); e pela “Nuvem Comunitária” somente 8 dos respondentes (4,6%) tiveram essa preferência.

O Gráfico 5 detalha as informações dos dados coletados.



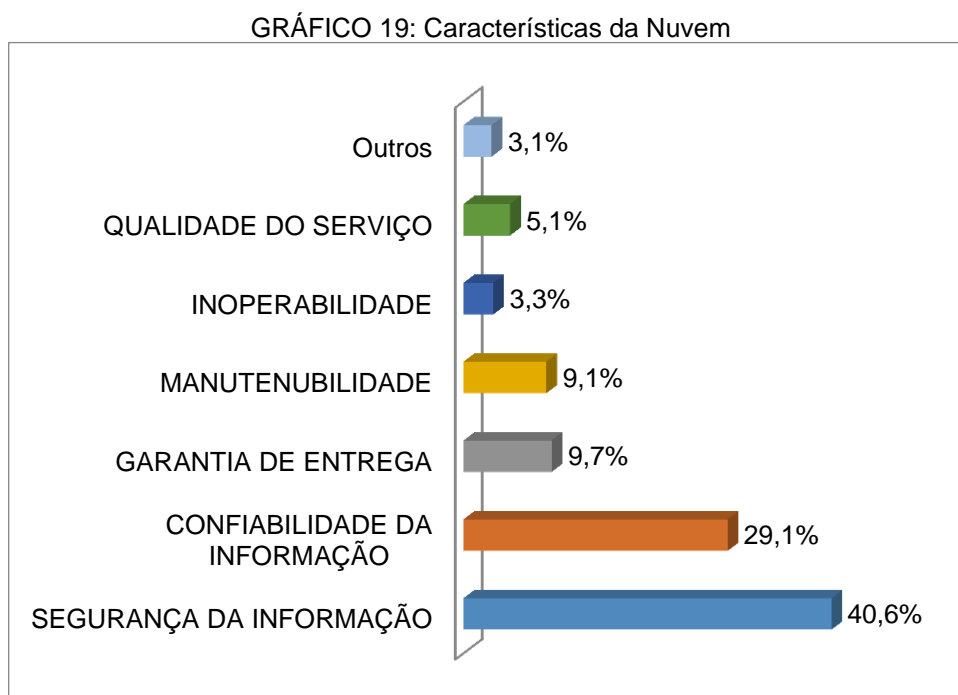
Fonte: próprio autor.

#### 4.2.2 Motivos específicos

Nesse quesito, a questão feita foi a seguinte: *qual seria o motivo específico para sua empresa não utilizar os serviços em Nuvem?*

Coletou-se o seguinte cenário: 71 respondentes (40.60%) não utilizam Nuvem por questões de “Segurança da Informação”; 51 respondentes (29.10%) informaram que a “Confiabilidade da informação” seria o principal motivo; 17 respondentes (9.70%) alegam que “Garantia de entrega do serviço” é o fator principal para essa não migração; para 16 respondentes (9.10%), a “Manutenibilidade” é o principal fator; 6 respondentes (3.3%) julgam a “Inoperabilidade” como o motivo maior para não aderir à arquitetura em nuvem; 9 respondentes (5.10%) declararam que é a “Qualidade do Serviço” que impossibilita essa adesão; e “Outros” obteve 5 respondentes (3,1%).

O Gráfico 6 detalha a informação coletada.



Fonte: próprio autor.

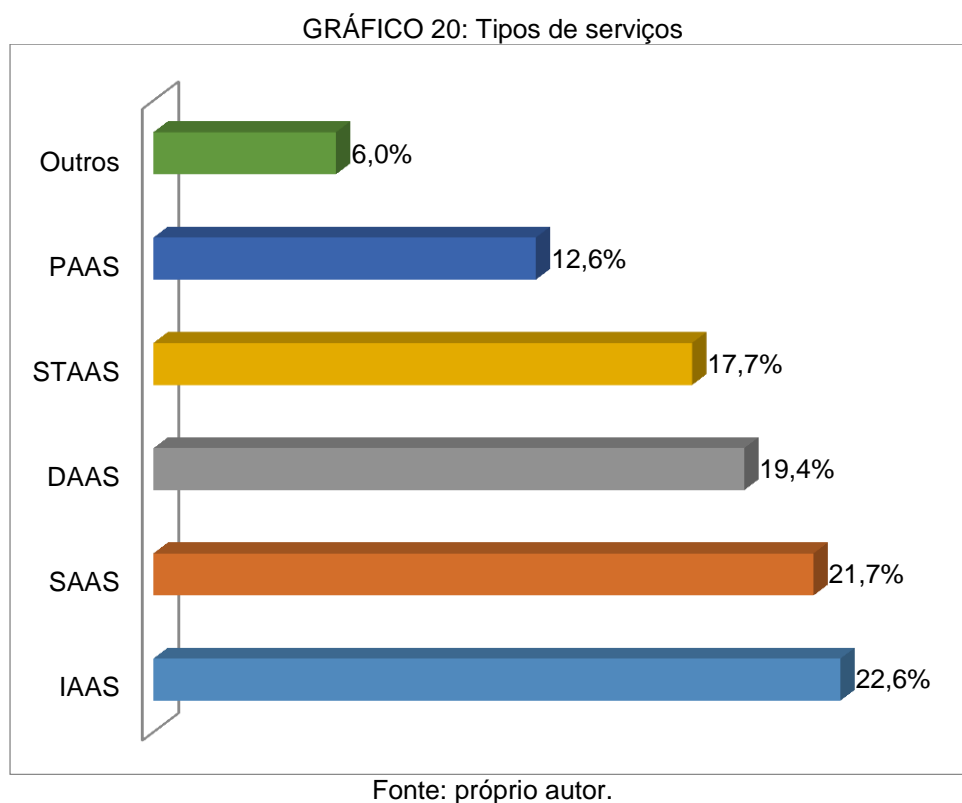
#### 4.2.4 Quanto à migração para um Serviço

Uma característica marcante da Nuvem é oferecer diversos ambientes de tecnologia como serviços, desde infraestrutura até mesmo funcionalidade com

especificidade bem peculiar. Para tal, realizou-se uma abordagem nesse sentido a partir da seguinte pergunta: *se sua Empresa/Organização resolvesse migrar para a Nuvem qual serviço utilizaria?*

A “Infraestrutura como Serviço” obteve 40 respondentes (22.80%); “Software como Serviço” obteve 38 respondentes (21.70%); 34 respondentes (19.40%) optaram por “Banco de dados como Serviço”; 31 respondentes (17.70%) preferiram “Armazenamento como Serviço”; “Plataforma como serviço” obteve 22 respostas (12.60%); “Forensic como Serviços” não obteve respostas; e “Outros” teve 10 respostas (0,6%).

O Gráfico 7 representa estas informações com mais clareza.



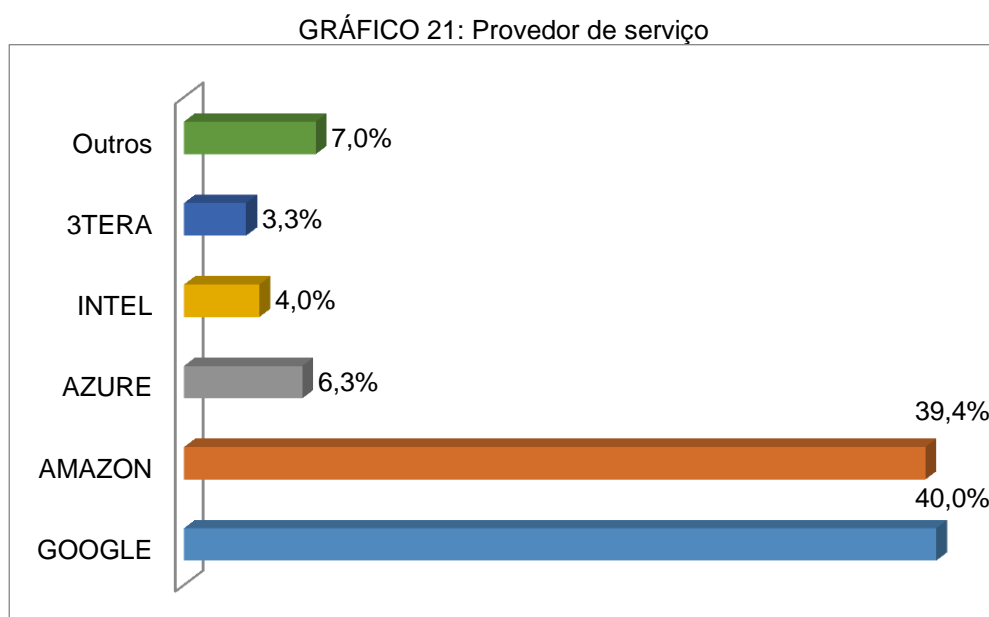
#### 4.2.5 Quanto ao provedor do Serviço

A Computação em Nuvem é uma grande aposta das empresas atualmente. Pensando nisso, empresas como *Google, Amazon, Azure, 3Tera, HP, Intel* e outras, que são fornecedoras de soluções em tecnologia, estão realizando investimentos cada vez maiores nesse segmento computacional, disponibilizando soluções nos mais diversificados tipos de serviços em Nuvem.

Sobre esse aspecto, a questão feita foi: *qual a empresa sua Empresa/Organização contrataria?*

A empresa *Google*, com 70 respondentes (40,0%), possui a preferência; a *Amazon* fica em segundo lugar com 69 respondentes (39,40%); a *Azure* da Microsoft tem a preferência de 11 respondentes (6,30%); a *Intel* tem 11 respondentes (4,00%); a *3Tera* obteve 4 respostas (3,40%); *Sun* e *HP* não obtiveram respostas; e “Outros” tiveram 12 respondentes (7,0%).

O Gráfico 8 define melhor essa coleta de dados.



Fonte: próprio autor.

### 4.3 Sob a perspectiva das Empresas ou Organizações que utilizam Computação em Nuvem

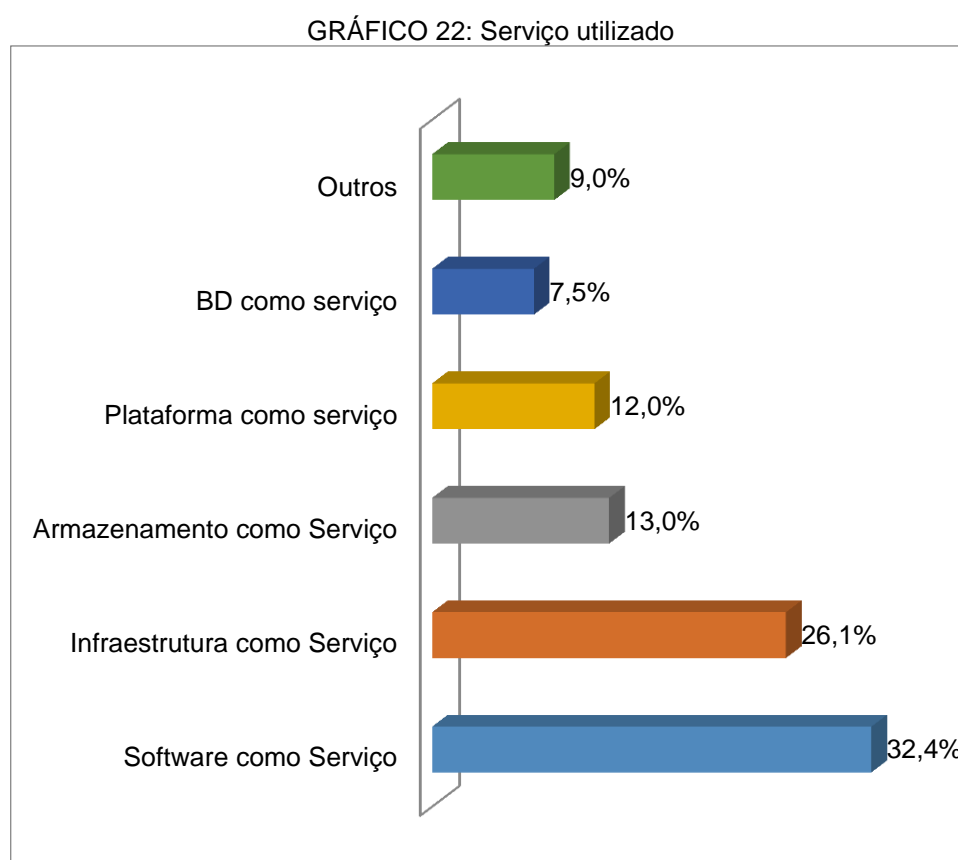
Para direcionar as respostas dos respondentes que utilizam algum tipo de serviço da Computação em Nuvem, desenvolveu-se um formulário específico. Com esse formulário, atingiu-se 92 respondentes. Cada questão desse formulário específico é tratada nas subseções a seguir.

#### 4.3.1 Quanto aos serviços utilizados

Neste questionamento, procurou-se identificar qual tipo de nuvem é utilizado. A pergunta foi: *qual serviço sua Empresa/Organização encontra-se utilizando?*

Obteve-se o seguinte resultado: dos 92 respondentes, 30 (32,40%) afirmaram utilizar “Software como Serviço”; “Infraestrutura como Serviço” representa a preferência de 24 dos respondentes (26,10%); “Armazenamento como Serviço” foi mencionado por 12 dos respondentes (13,00%); “Plataforma como Serviço” representa a preferência de 11 dos respondentes (12,00%); “Banco de dados como Serviço” foi a preferência de 6 dos respondentes (7,50%); “Forensic como Serviço” não obteve resultados; e “Outros” foi respondido por 9 participantes (9,0%).

Os dados estão representados no Gráfico 9.



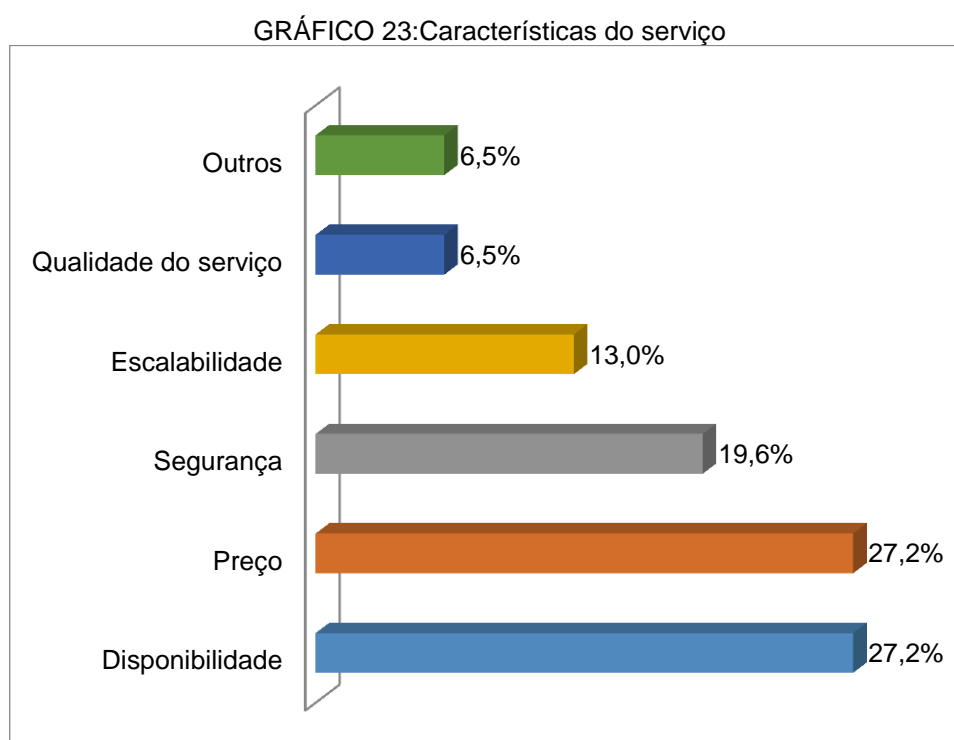
Fonte: próprio autor.

#### 4.3.2 Quanto às características (Serviços utilizados)

Também se optou pela identificação do grau de satisfação dos contratantes em relação a alguma característica da Nuvem utilizada no momento. A indagação feita foi: *das características abaixo indique qual a sua Empresa/Organização optou pela nuvem?*

Coletaram-se as seguintes respostas: “Disponibilidade” é representada por 25 dos respondentes (27,20%) – a mesma quantidade se verificou em relação ao “Preço”; “Segurança” é representada por 18 dos respondentes (19,60%); “Escalabilidade” foi a opção de 12 dos respondentes (13,00%); “Qualidade dos Serviços” apresenta 6 respondentes (6,5%); e “Outros” foi a opção escolhida por 6 dos respondentes (6,5%).

Esta coleta de dados pode ser visualizada no Gráfico 10.



Fonte: próprio autor.

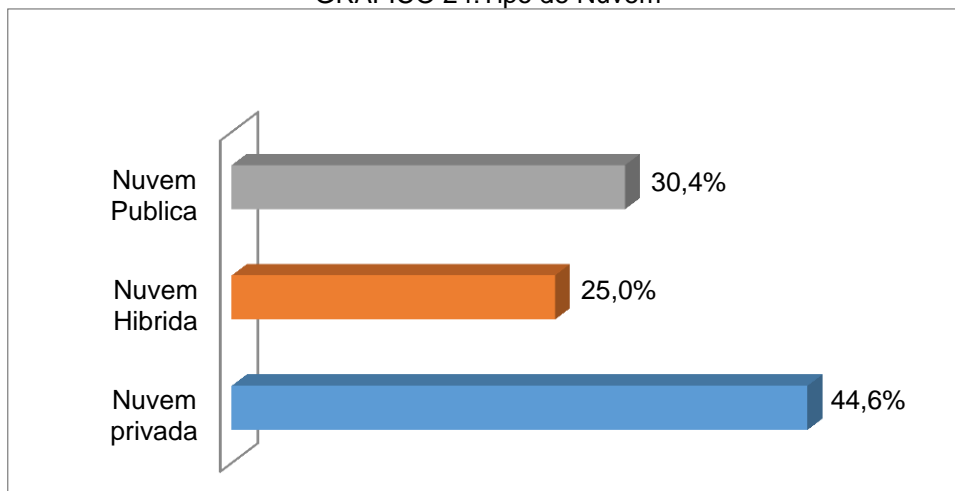
#### 4.3.3 Quanto ao tipo de nuvem utilizada

Nessa questão identificou-se qual o tipo de nuvem que os contratantes estão utilizando no momento. A pergunta feita foi: *que tipo de nuvem sua Empresa/Organização utiliza?*

O resultado dessa questão foram os seguintes: a “Nuvem Privada” foi a opção de 41 dos respondentes (44,60%); 28 dos respondentes (30,40%) preferem a “Nuvem Pública”; e a “Nuvem Híbrida” foi escolhida por 23 dos respondentes (25,00%). Nenhum respondente optou por “Nuvem Comunitária”.

Para representar essa coleta temos o Gráfico 11.

GRÁFICO 24: Tipo de Nuvem



Fonte: próprio autor.

#### 4.3.4 Segurança da Nuvem

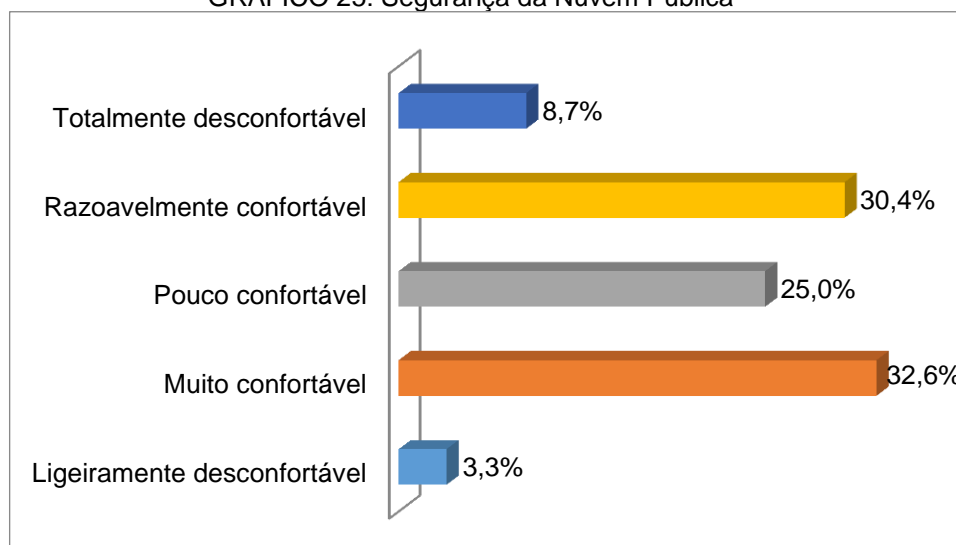
A segurança do ambiente computacional possui seus questionamentos e, com a Computação em Nuvem, é provável que este seja um dos principais temas discutidos entre os contratantes e investidores. Para identificar qual o nível de satisfação das empresas em relação à segurança, questionou-se: *em caso de Nuvem Pública, sua Empresa/Organização sente-se confortável quanto à segurança na sua utilização?*

Verificou-se o seguinte resultado: 30 dos respondentes (32,60%) afirmaram estar “Muito Confortável” com a segurança do serviço contratado; 28 dos respondentes (30,40%) informaram estar “Razoavelmente Confortável”; 8 dos respondentes (8,70%) admitiram estar “Totalmente Desconfortável” em relação à segurança da Computação em Nuvem; 23 respondentes (25,00%) selecionaram a opção “Poucos Confortável”; e 3 (3,30%) “Ligeiramente Confortável”.

Para representar estes dados, apresenta-se o Gráfico 12.



GRÁFICO 25: Segurança da Nuvem Pública



Fonte: próprio autor.

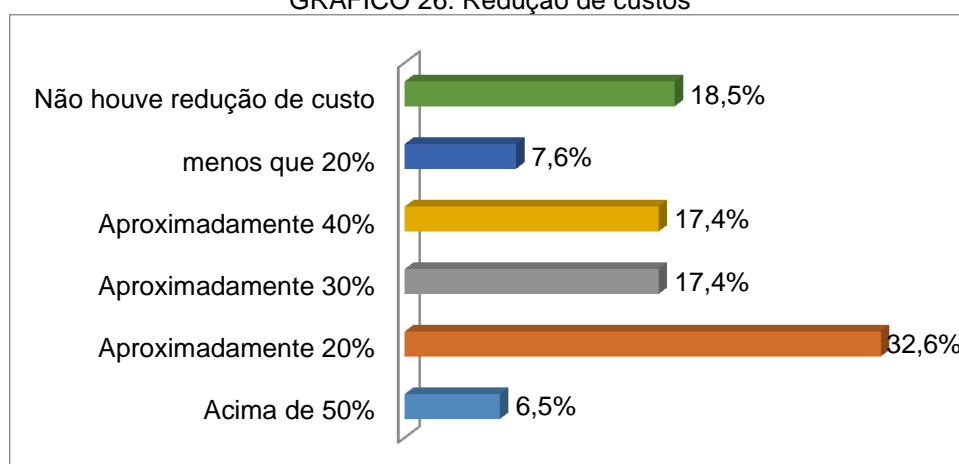
#### 4.3.5 Quanto à redução de custo

Existe um apelo forte da Computação em Nuvem que é a redução do custo nos investimentos. Sobre essa questão, o propósito é identificar se as empresas que aderiram realmente alcançaram essa redução em seus custos. A indagação feita foi: *qual o percentual de redução de custo que sua empresa alcançou com a migração para a Nuvem?*

Como resultado, constatou-se o que segue: 30 dos respondentes (32,60%) disseram ter reduzido “Aproximadamente 20%” nos custos; 17 dos respondentes (18,50%) afirmaram que “Não houve redução” de custo com a contratação da Nuvem; 16 dos respondentes (17,40%) afirmaram ter reduzido “Aproximadamente 30%” dos custos com o investimento da Computação em Nuvem; 16 dos respondentes (17,40%) afirmaram ter reduzido “Aproximadamente 40%”; 7 dos respondentes (7,6%) afirmam ter reduzido “Menos que 20%” com a Nuvem; 6 dos respondentes (6,50%) consideram que houve uma redução considerável de “Aproximadamente 50%” dos investimentos.

Esses dados relacionados ao custo podem ser visualizados no Gráfico 13.

GRÁFICO 26: Redução de custos



Fonte: próprio autor.

#### 4.4 Resumo do Capítulo

Neste capítulo apresentou-se a coleta dos dados, a partir de um total de 14 (quatorze) questões para esta pesquisa investigativa, cujas informações são pré-requisitos para o capítulo seguinte de análise de dados, em que se realiza um estudo comparativo entre as respostas com o objetivo de encontrar correlação, possibilitando a identificação dos melhores casos a partir do treinamento realizado. A distribuição dos gráficos foi apresentada de acordo com a perspectiva dos formulários (dados preliminares, não utilização de Nuvem e utilização de Computação em Nuvem). No capítulo seguinte, aborda-se a análise dos resultados explicitados nesta seção, aplicando aprendizado automático e fazendo uso da tarefa de classificação com algoritmo inteligente J48 (Árvore de Decisão).

## 5 ANÁLISE DOS DADOS

O capítulo de análise dos dados tem como objetivo encontrar as respostas para os questionamentos realizados na pesquisa. A partir dessas análises, são apresentadas duas hipóteses como respostas para esta pesquisa.

Na seção 5.1 apresenta-se a extração do conhecimento através dos conjuntos de treinamentos com os respondentes que não usam Nuvem e com aqueles que utilizam algum serviço Nuvem; na seção 5.2, apresentam-se as discussões dos resultados, apresentando os melhores casos descritos em cada hipótese; e, por fim, na seção 5.3, o resumo do capítulo.

Nesta etapa de análise, a ferramenta utilizada foi a WEKA 3.6 (ferramenta de aprendizado automático), que foi fundamental no processo de aprendizado e construção do conhecimento. Criou-se um conjunto de treinamento baseado nas respostas dos participantes com as instâncias de quem usa ou não a Computação em Nuvem. A partir desse conjunto de treinamento, realizou-se o processamento dos dados visando à descoberta de conhecimento através do aprendizado, utilizando-se para isso algoritmos inteligentes da área de Inteligência Artificial, especificamente o aprendizado automático. Nos testes, utilizaram-se algumas técnicas existentes na ferramenta WEKA, com recursos de aprendizagem como: classificação, clusterização e associação, tendo disponível, para cada técnica, diversos algoritmos inteligentes. Com base nesses recursos, optou-se pela utilização dessa ferramenta com o propósito de encontrar a melhor técnica e, conseqüentemente, o melhor resultado no aprendizado.

### 5.1 Treinamentos

Nos treinamentos, a partir dos atributos quantitativos, foi utilizada a hierarquia do aprendizado supervisionado, aplicando a tarefa de classificação com árvore de decisão, mediante a utilização do algoritmo de aprendizado J48. Isso permitiu a classificação das respostas e geração da árvore de decisão como resultado, aplicando-se o processamento “com poda”, na qual o conhecimento adquirido é representado por meio de regras. Assim os percentuais dos “Acertos” e “Erros” (intervalo de confiança) visam “a obtenção de uma regra de decisão final que dá mais importância à classe de interesse” (CASTRO e BRAGA, 2011, p. 14) alcançada

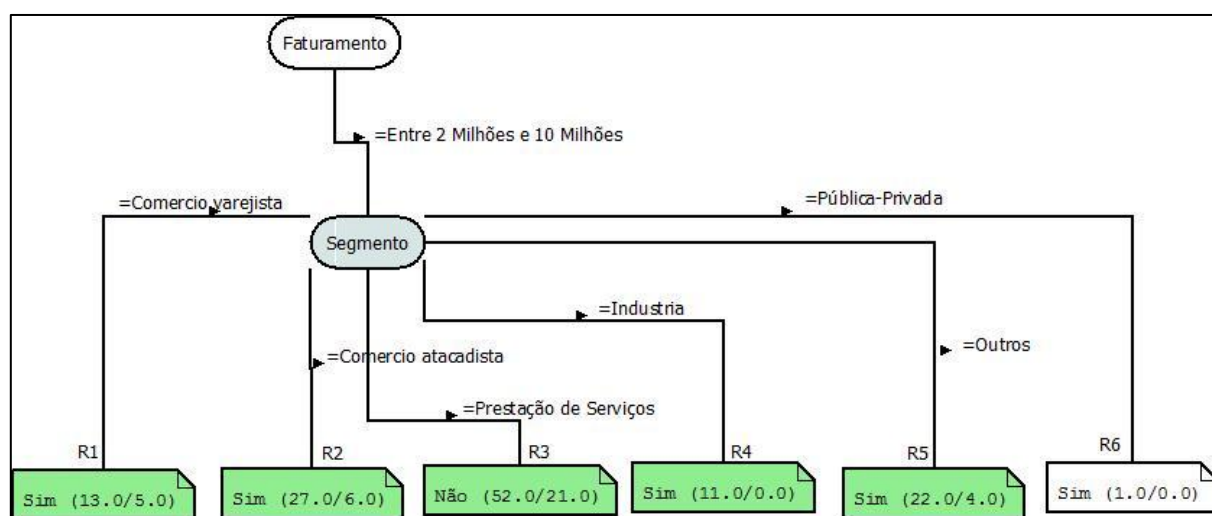
pelo treinamento, em que o “acerto” representa a quantidade de instâncias classificadas (treinadas) de forma correta pelo classificador, sendo que o contrário acontece com os “erros”, em que o classificador representa as instâncias classificadas de forma incorretamente, ou seja, para o erro, o classificador realiza a tentativa de classificação, porém sem êxito. Para facilitar a identificação do resultado apresentado nas subárvores, as folhas foram identificadas através da letra “R” (de ‘resultado’), e a partir da árvore principal foram selecionados os melhores casos, conforme o exemplo (R1, R2, R3 ... Rn).

### 5.1.1 Treinamento para identificar os indícios para contratar ou não

O primeiro conjunto de treinamento foi projetado no intuito de identificar os participantes que contratariam ou não a Nuvem. Pode-se ter uma visão geral da árvore de decisão construída na figura presente no Apêndice B1 deste trabalho. Nesse processamento foi gerada uma árvore de tamanho 12 (doze) níveis ou nós e com 54 (cinquenta e quatro) folhas. Para a visualização em um nível mais detalhado, foram destacados da árvore gerada os principais ramos, a partir da classificação dos resultados para os vetores de atributos.

Inicia-se com a subárvore cujos nós são Faturamento e Segmento. O classificador identificou os melhores resultados, que podem ser visualizados na Figura 1, onde os resultados são representados de R1 a R6.

FIGURA 11: Árvore de decisão – Faturamento entre 2 Milhões e 10 Milhões

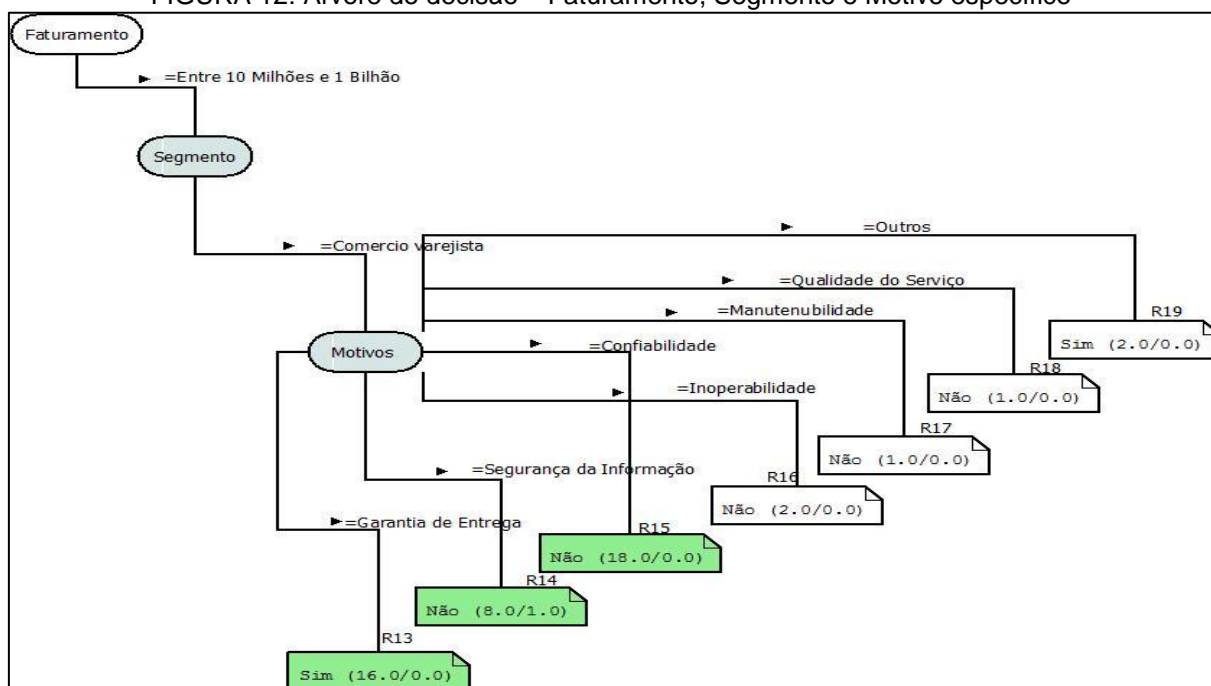


Fonte: próprio autor.

- (R1) Empresas do segmento de “Comércio Varejista” cujo faturamento anual se encontra “Entre: 2 milhões e 10 milhões”. O treinamento classificou que contrariariam Nuvem, com base nas 13 (treze) ocorrências do treinamento, sendo identificados 38,4% com 5 (cinco) instâncias classificadas incorretamente.
- (R2) Empresas do segmento de “Comércio Atacadista” cujo faturamento anual se encontra “Entre: 2 milhões e 10 milhões”. O treinamento classificou que utilizariam Nuvem, com base nas 27 (vinte e sete) ocorrências do treinamento, sendo identificados 22,22% com 6 (seis) instâncias classificadas incorretamente.
- (R3) Empresas do segmento de “Prestação de Serviços” e com faturamento anual se encontra “Entre: 2 milhões e 10 milhões”. O treinamento classificou que a maioria não utilizaria Nuvem, com base nas 52 (cinquenta e duas) ocorrências do treinamento, sendo identificados 40,38% e 21 (vinte e uma) instâncias classificadas incorretamente.
- (R4) Empresas do segmento “Industrial” cujo faturamento anual se encontra “Entre: 2 milhões e 10 milhões”. O treinamento classificou que utilizariam Nuvem, com base nas 11 (onze) ocorrências do treinamento e 100% de acertos, o treinamento foi bem-sucedido para essa classificação.
- (R5) As empresas que atuam em “OUTROS segmentos” com faturamento anual “Entre: 2 milhões e 10 milhões”. O treinamento classificou que utilizariam Nuvem, com base nas 22 (vinte e duas) ocorrências do treinamento, sendo identificados 18,18% com 4 (quatro) instâncias classificadas incorretamente.
- Vale ressaltar que, para essa ramificação, a folha R6 foi considerada insignificante em virtude das ocorrências desse treinamento.

Para a subárvore cujos nós são Faturamento, Segmento e Motivo específico para não usar Computação em Nuvem, o classificador identificou os melhores casos, que podem ser visualizados melhor na Figura 2, onde os resultados são representados de R13 a R19.

FIGURA 12: Árvore de decisão – Faturamento, Segmento e Motivo específico

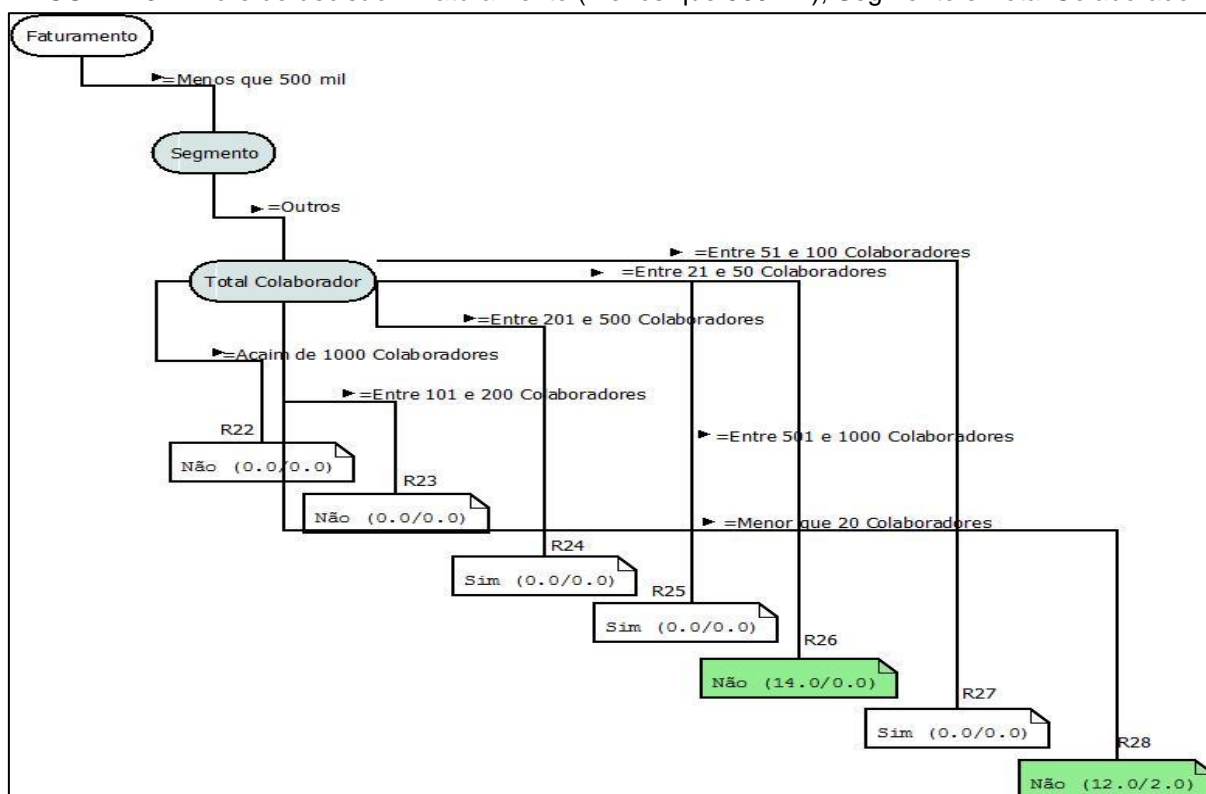


Fonte: próprio autor.

- (R13) Empresas cujo faturamento anual se encontra “Entre: 10 milhões e 1 Bilhão” e segmento de “Comércio Varejista” que possuem como motivo específico “Garantia de Entrega”. O treinamento classificou que contratariam a Nuvem, com base nas 15 (quinze) ocorrências do treinamento e 100% de acertos para essa classificação.
- (R14) Empresas cujo faturamento anual se encontra “Entre: 10 milhões e 1 Bilhão” e segmento de “Comércio Varejista” que possuem como motivo específico “Segurança da Informação”. O treinamento classificou que a maioria não contrataria a Nuvem, com base nas 8 (oito) ocorrências do treinamento, sendo identificadas 12,5%, com 1 (uma) instância classificada incorretamente.
- (R15) Empresas cujo faturamento anual se encontra “Entre: 10 milhões e 1 Bilhão” e segmento de “Comércio Varejista” que possuem como motivo específico como “Confiabilidade”. O treinamento classificou que contratariam Nuvem com base nas 18 (dezoito) ocorrências do treinamento, sendo identificadas 100% das instâncias foram classificadas corretamente.
- Para as classificações dos resultados R16, R17, R18 e R19, o classificador considerou os resultados insignificantes em virtude das ocorrências desse treinamento.

Para a subárvore em que os nós são Faturamento, Segmento e Total Colaborador, o classificador identificou os melhores resultados que podem ser analisados na Figura 3, onde os resultados são representados de R22 a R28.

FIGURA 13: Árvore de decisão – Faturamento (Menos que 500 mil), Segmento e Total Colaborador



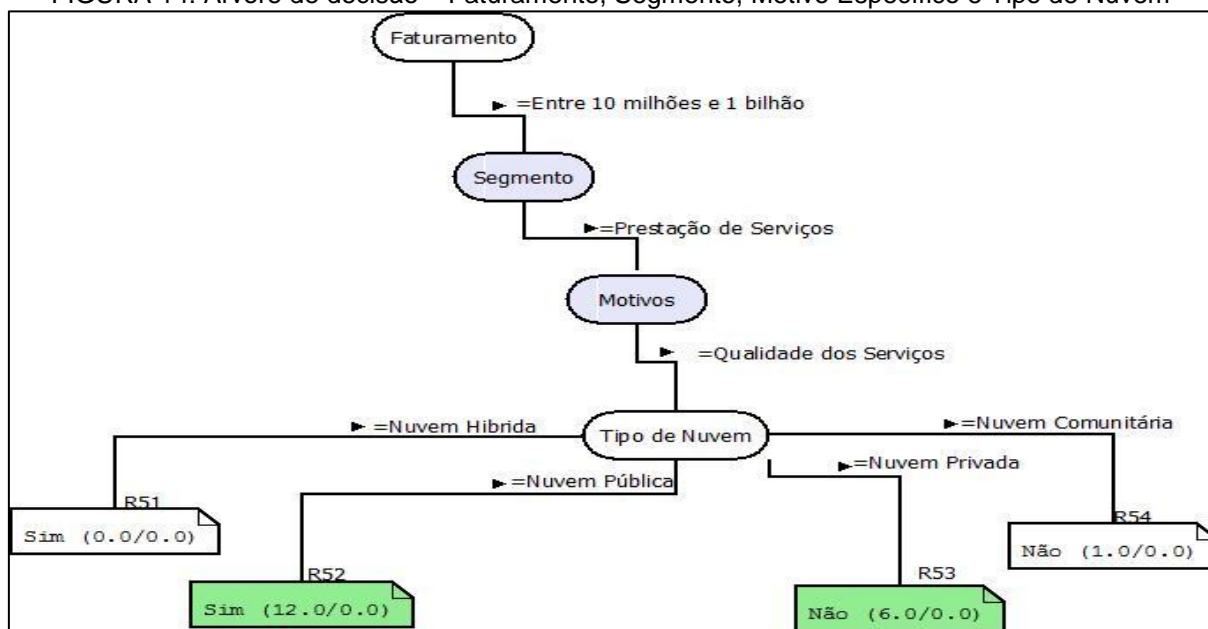
Fonte: próprio autor.

- (R26) Empresas cujo faturamento anual é “Menor que 500 mil” e segmento “Outros” e total de colaboradores “Entre: 21 e 50 colaboradores”. O treinamento classificou que 100% não contrataria a Nuvem, com base nas 14 (quatorze) ocorrências do treinamento para essa classificação.
- (R28) Empresas cujo faturamento anual é “Menor que 500 mil” e segmento “Outros” e total de colaboradores “Menor que 20 colaboradores”. O treinamento classificou que não contratariam a Nuvem, com base nas 12 (doze) ocorrências do treinamento, sendo identificadas que 16,66% ou 2 (duas) instâncias foram classificadas incorretamente.
- Os resultados R22, R23, R24, R25 e R27 foram considerados pelo classificador insignificantes em virtude das ocorrências desse treinamento.

Para a subárvore com nós Faturamentos, Segmento de atuação, Motivo específico e Tipo de Nuvem, o classificador identificou os dois melhores casos, que

são detalhados melhor na Figura 4, onde os resultados são representados de R51 a R54.

FIGURA 14: Árvore de decisão – Faturamento, Segmento, Motivo Específico e Tipo de Nuvem



Fonte: próprio autor.

- (R52) Empresas cujo faturamento anual está “Entre 10 milhões e 1 bilhão” e são do segmento de “Prestação de Serviços” e têm como motivo específico para não contratar a Nuvem a “Qualidade dos Serviços” e tipo de nuvem a “Nuvem Pública”. O treinamento classificou que contratariam Nuvem, com base nas 12 (doze) ocorrências do treinamento e 100% de acerto para essa classificação.
- (R53) Empresas cujo faturamento anual está “Entre 10 milhões e 1 bilhão” e são do segmento de “Prestação de Serviços” e têm como motivo específico para não contratar a Nuvem a “Qualidade dos Serviços” e tipo de nuvem a “Nuvem Privada”. O treinamento classificou que a maioria não utilizaria Nuvem, com base nas 6 (seis) ocorrências realizadas pelo treinamento e 100% das instâncias classificadas incorretamente.
- Os resultados R51 e R54 não foram significativos.



### 5.1.1.1 Matriz de Confusão – Treinamento Contrataria ou não Nuvem

A matriz de confusão quantifica quantos exemplos da base de dados utilizada seriam classificados corretamente pelo modelo construído e representado na diagonal principal, sendo que os que seriam classificados incorretamente são representados pela diagonal transversal (FONSECA, 1994).

A importância da matriz de confusão é poder mostrar como o modelo de classificação se comporta com os dados obtidos através dos questionamentos, quantos casos ele acerta/erra, quantos são os falsos positivos. Também pode-se dizer que é um resultado geral das instâncias classificadas correta ou incorretamente.

Essa interpretabilidade pode ser vista na Tabela 1 com o resultado do conjunto de treinamento dos respondentes que não usam nuvem quando configurado com poda. O algoritmo classificador J48 apresentou os seguintes resultados na matriz de confusão:

TABELA 4: Matriz de Confusão – Treinamento Contrataria ou não nuvem

A	B	
147	32	Não
29	59	Sim

Fonte: próprio autor.

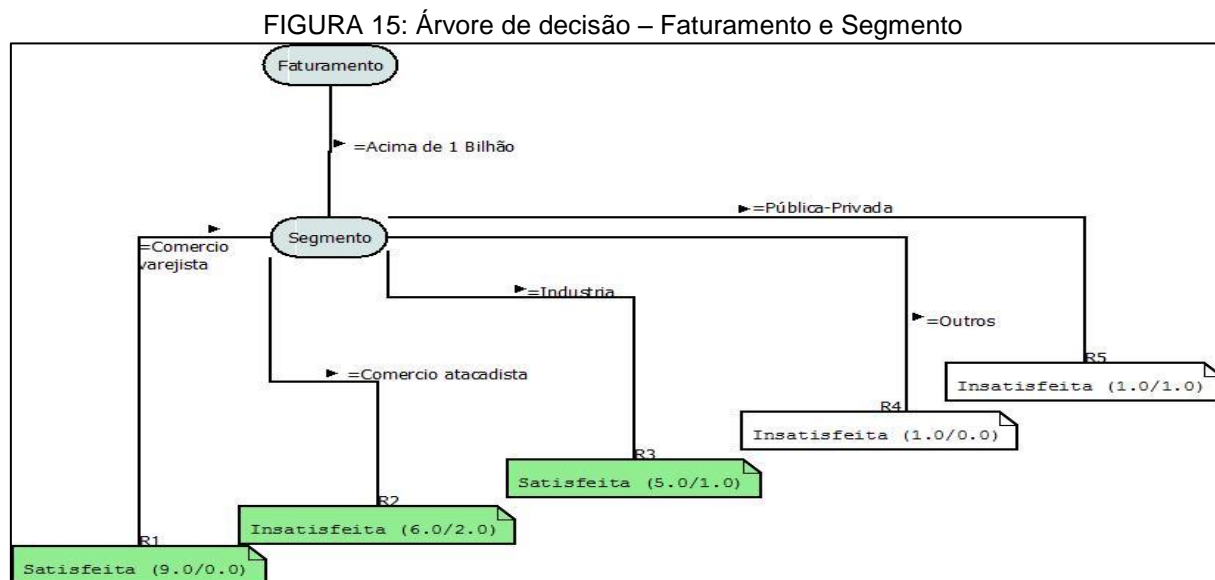
Na interpretabilidade da matriz temos: na linha diagonal principal, representada pelos valores 147 (cento e quarenta e sete) e 59 (cinquenta e nove), que somados totalizam 206 (duzentos e seis) instâncias classificadas corretamente; a mesma operação é realizada para a diagonal transversal, representada pelos valores 29 (vinte e nove) e 32 (trinta e dois), somando 61 (sessenta e uma) instâncias classificadas incorretamente, gerando um total geral de 267 (duzentos e sessenta e sete) instâncias para ambas as diagonais.

- ✓ Instâncias classificadas corretamente com porcentagem:  $(206/267)*100=77,1535\%$ .
- ✓ Instâncias classificadas incorretamente tiveram a porcentagem:  $(61/267)*100=22,8464\%$  de acurácia.

Em geral, pode-se concluir que o "erro médio" seria de 22,8464%. Assim, prevê-se que, quando se utilizar o modelo para classificar com poda, erra-se em 22,8464% dos casos.

### 5.1.2 Treinamento para identificar o motivo da Satisfação ou Insatisfação

Para o treinamento onde a linha de investigação é identificar se as empresas que estão utilizando nuvem no momento estão satisfeitas ou insatisfeitas com os serviços utilizados, foi desenvolvido um conjunto de treinamento para a averiguação dessas hipóteses. Para o processamento desse, permaneceu-se a configuração com poda, que resultou em uma árvore de tamanho 30 (trinta) folhas com 7 (sete) níveis. Nessa árvore, destacaram-se seus ramos de acordo com as melhores classificações. A visão geral da árvore de decisão se encontra no apêndice B2 desta dissertação. Para a subárvore cujo nó principal é Faturamento e o nó secundário é Segmento, o classificador identificou os melhores casos, detalhados na Figura 5, onde os resultados são representados de R1 a R5.



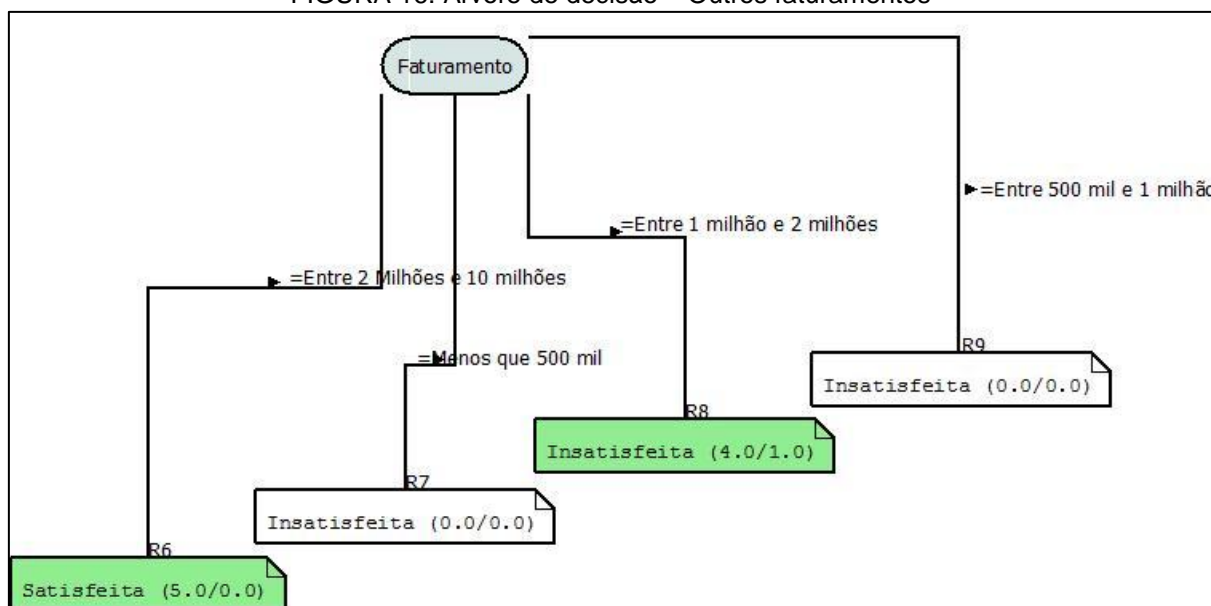
Fonte: próprio autor.

- (R1) Empresas do segmento de “Comércio Varejista” cujo faturamento se encontra acima de 1 Bilhão anual. O treinamento classificou que estão satisfeitas com contratação da Nuvem, com base nas 9 (nove) ocorrências do treinamento e 100% de acertos para essa classificação.

- (R2) Empresas do segmento de “Comércio Atacadista” cujo faturamento é superior a 1 bilhão anual. O treinamento classificou que estão insatisfeitas com a Nuvem, com base nas 6 (seis) ocorrências do treinamento, sendo identificadas 33,00% das instâncias classificadas incorretamente.
- (R3) Empresas do segmento “Industrial” cujo faturamento é “Acima de 1 Bilhão” anual, o treinamento classificou que estão satisfeitas com a Nuvem com base nas 5 (cinco) ocorrências do treinamento, sendo identificado que 20% das instâncias foram classificadas incorretamente.
- Quanto às classificações dos resultados R4, e R5, o classificador considerou os resultados insignificantes em virtude das ocorrências desse treinamento.

Para a subárvore em que os faturamentos são “menos que 500 mil”, “Entre: 500 e 1 milhão”, “Entre: 1 milhão e 2 Milhões” e “Entre: 2 milhões e 10 milhões” anuais, o classificador identificou os melhores resultados que podem ser analisados na Figura 6, onde os resultados são representados de R6 a R9.

FIGURA 16: Árvore de decisão – Outros faturamentos



Fonte: próprio autor.

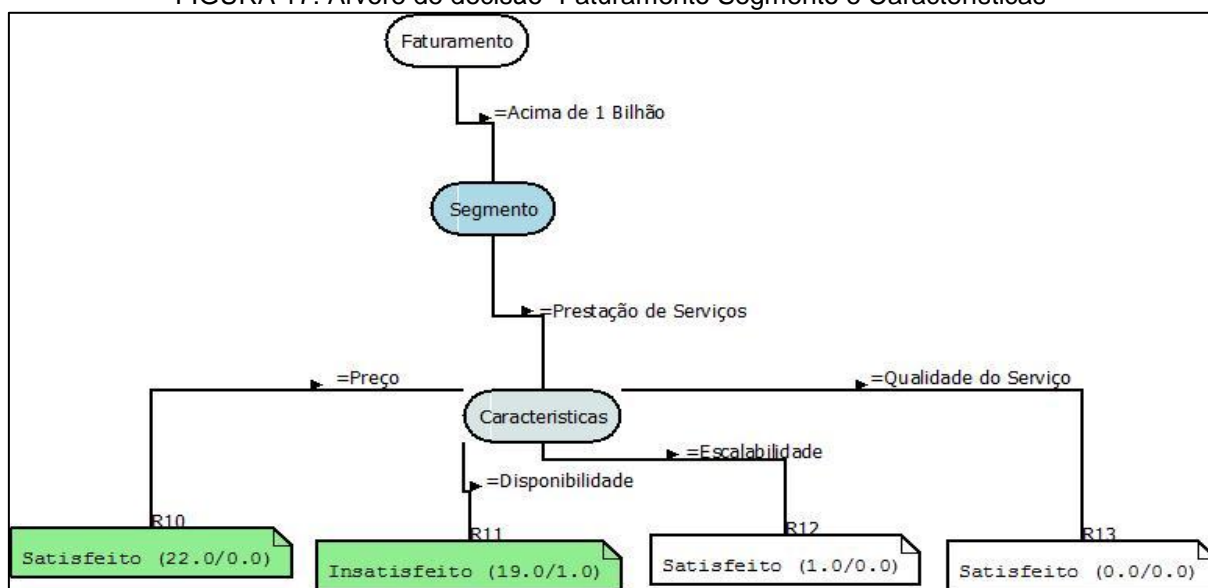
- (R6) Empresas com faturamento “Entre: 2 milhões e 10 milhões” anual, o treinamento classificou que estão satisfeitas com a contratação da Nuvem

com base nas 5 (cinco) ocorrências do treinamento, sendo identificado que 100% das instâncias foram classificadas corretamente.

- (R8) Empresas com faturamento “Entre: 2 milhões e 10 milhões” anual, o treinamento classificou que estão insatisfeitas com a contratação da Nuvem com base nas 4 (quatro) ocorrências do treinamento, sendo identificado que 25% das instâncias foram classificadas incorretamente.
- No que diz respeito às classificações dos resultados R7 e R9, o classificador considerou os resultados insignificantes em virtude das ocorrências desse treinamento.

Para a subárvore cujos nós são Faturamento, Segurança da Nuvem Pública e Total colaborador, os melhores resultados são apresentados na Figura 7, de R26 a R30.

FIGURA 17: Árvore de decisão- Faturamento Segmento e Características



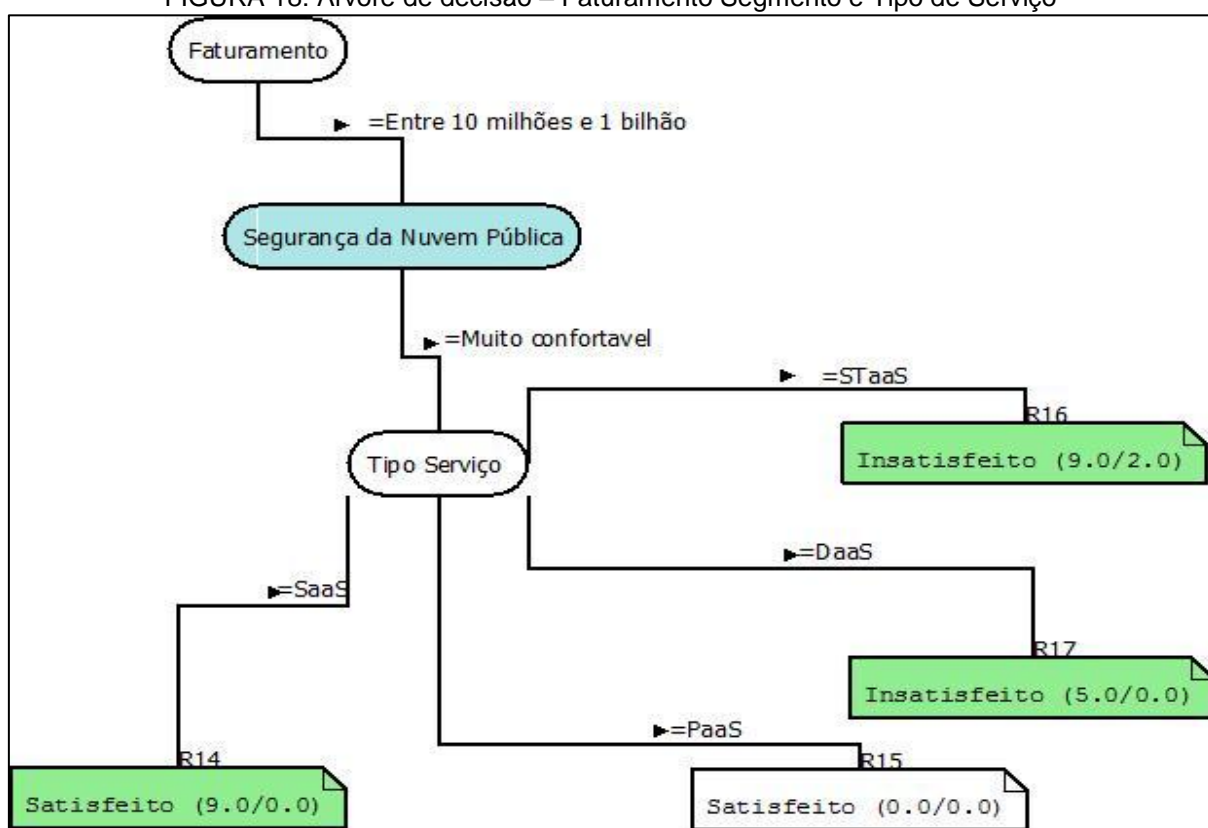
Fonte: próprio autor.

- (R10) Empresas cujo faturamento anual é “Acima de 1 bilhão” e têm como “Segmento” “Prestação de Serviços” e que possuem “Características” de “Preço”. O treinamento classificou que estão satisfeitas com a contratação da Nuvem com base nas 22 (vinte e duas) ocorrências do treinamento, sendo identificado que 100% das instâncias foram classificadas corretamente.

- (R11) Empresas cujo faturamento anual é “Acima de 1 bilhão”, do “Segmento” “Prestação de Serviços” e que possuem “Característica” de “Disponibilidade”. O treinamento classificou que estão insatisfeitas com a contratação da Nuvem com base nas 19 (dezenove) ocorrências do treinamento, sendo identificado que 5,2% das instâncias foram classificadas incorretamente.
- Quanto à classificação dos resultados R12 e R13, o classificador considerou insignificante em virtude das ocorrências desse treinamento.

Para a subárvore cujos nós são Faturamento, Segmento de atuação e Tipo de serviços utilizados, os resultados são representados de R14 a R17 na Figura 8.

FIGURA 18: Árvore de decisão – Faturamento Segmento e Tipo de Serviço



Fonte: próprio autor.

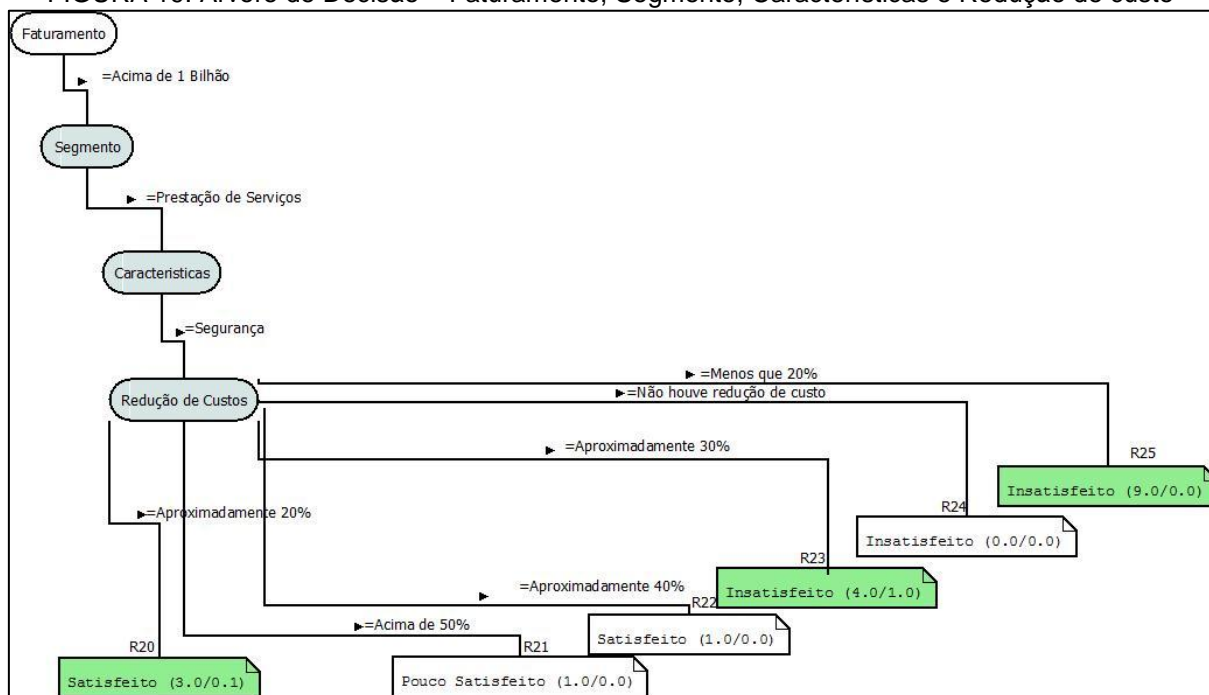
- (R14) Empresas cujo faturamento anual é “Entre: 10 milhões e 1 bilhão”, e que estão “muito confortáveis” com a “Segurança da Nuvem Pública” e o “Tipo de serviços” utilizado é “SaaS”. O treinamento classificou que estão satisfeitas com a contratação da Nuvem com base nas 9 (nove) ocorrências

do treinamento, sendo identificado que 100% das instâncias foram classificadas corretamente.

- (R16) Empresas com faturamento anual “Entre: 10 milhões e 1 bilhão”, e que estão “muito confortáveis” com a “Segurança da Nuvem Pública” e o “Tipo de serviços” utilizado é “STaaS”. O treinamento classificou que estão insatisfeitas com a contratação da Nuvem com base nas 9 (nove) ocorrências do treinamento, sendo identificado que 22,22% das instâncias foram classificadas incorretamente.
- (R17) Empresas com faturamento anual “Entre: 10 milhões e 1 bilhão”, e que estão “muito confortáveis” com a “Segurança da Nuvem Pública” e o “Tipo de serviços” utilizado é “DaaS”. O treinamento classificou que estão satisfeitas com a contratação da Nuvem com base nas 5 (cinco) ocorrências do treinamento, sendo identificado que 100% das instâncias foram classificadas corretamente.
- Quanto à classificação do resultado R15, o classificador considerou insignificante em virtude das ocorrências desse treinamento.

Para a subárvore em que os nós são Faturamento, Segmento de atuação, características e Redução de Custos, os resultados são representados de R20 a R25 na Figura 9.

FIGURA 19: Árvore de Decisão – Faturamento, Segmento, Características e Redução de custo



Fonte: próprio autor.

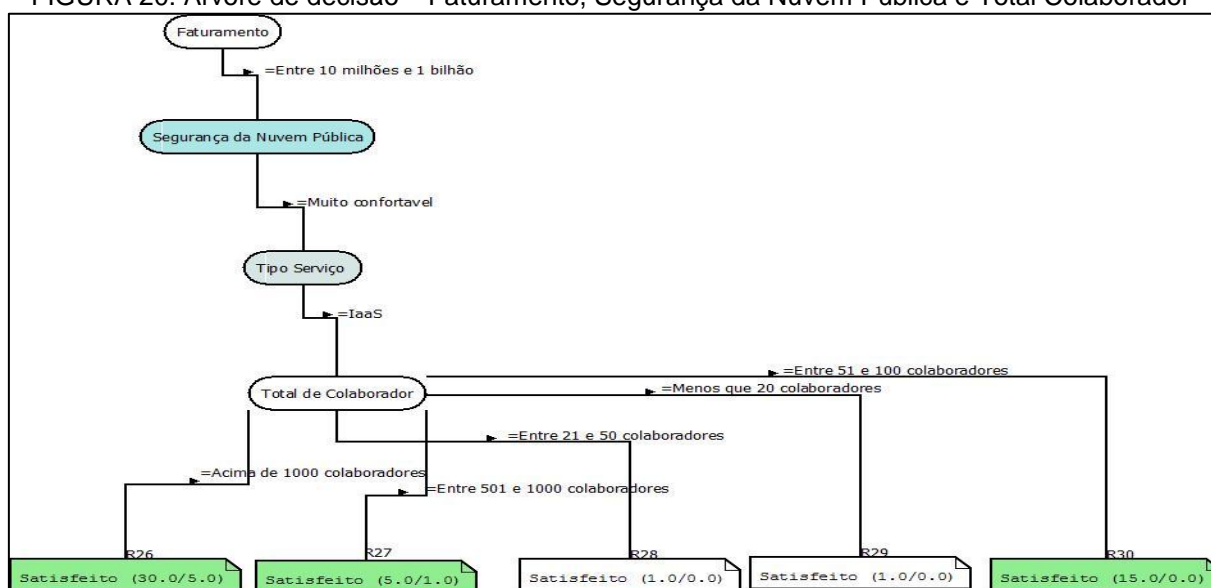
- (R20) Empresas cujo faturamento anual é “Acima de 1 bilhão”, e são do “Segmento” da área de “Prestação de Serviços” e que possuem como “Características” “Segurança” e tiveram “Redução de Custos” “Aproximadamente 20%”. O treinamento classificou que estão satisfeitas com a contratação da Nuvem com base nas 3 (três) ocorrências do treinamento, sendo identificado que 33,33% das instâncias foram classificadas incorretamente.
- (R23) Empresas cujo faturamento anual é “Acima de 1 bilhão”, e são do “Segmento” da área de “Prestação de Serviços” e que possuem como “Características” “Segurança” e tiveram “Redução de Custos” “Aproximadamente 30%”. O treinamento classificou que estão insatisfeitas com a contratação da Nuvem com base nas 4 (quatro) ocorrências do treinamento, sendo identificado que 25% das instâncias foram classificadas incorretamente.
- (R25) Empresas cujo faturamento anual é “Acima de 1 bilhão”, e são do “Segmento” da área de “Prestação de Serviços” e que possuem como “Características” “Segurança” e tiveram “Redução de Custos” “Menos que 20%”. O treinamento classificou que estão insatisfeitas com a contratação da

Nuvem com base nas 9 (nove) ocorrências do treinamento, sendo identificado que 100% das instâncias foram classificadas corretamente.

- Quanto à classificação dos resultados R21, R22 e R24, o classificador considerou insignificante em virtude das ocorrências desse treinamento.

Para a subárvore de nós Faturamento, Segurança da Nuvem Pública e Total colaborador os resultados são representados de R26 a R30 na figura 10.

FIGURA 20: Árvore de decisão – Faturamento, Segurança da Nuvem Pública e Total Colaborador



Fonte: próprio autor.

- (R26) Empresas cujo faturamento anual é “Entre: 10 milhões e 1 bilhão”, e que estão “muito confortáveis” com a “Segurança da Nuvem Pública” e o “Tipo de serviços” utilizado é “IaaS” e que possuem um “Total colaborador” “Acima de 1000 colaboradores”. O treinamento classificou que estão satisfeitas com a contratação da nuvem com base nas 30 (trinta) ocorrências do treinamento, sendo identificado que 16,66% das instâncias foram classificadas incorretamente.
- (R27) Empresas cujo faturamento anual é “Entre: 10 milhões e 1 bilhão”, e que estão “muito confortáveis” com a “Segurança da Nuvem Pública” e o “Tipo de serviços” utilizado é “IaaS” e que possuem um “Total colaborador” “Entre: 500 e 1000 colaboradores”. O treinamento classificou que estão satisfeitas com a contratação da nuvem com base nas 5 (cinco) ocorrências



do treinamento, sendo identificado que 20% das instâncias foram classificadas incorretamente.

- (R30) Empresas cujo faturamento anual é “Entre: 10 milhões e 1 bilhão”, e que estão “muito confortáveis” com a “Segurança da Nuvem Pública” e o “Tipo de serviços” utilizado é “IaaS” e que possuem um “Total colaborador” “Entre: 51 e 100 colaboradores”. O treinamento classificou que estão satisfeitas com a contratação da Nuvem com base nas 5 (cinco) ocorrências do treinamento, sendo identificado que 20% das instâncias foram classificadas incorretamente.
- Quanto à classificação dos resultados R28 e R29, o classificador considerou insignificante em virtude das ocorrências desse treinamento.

#### 5.1.2.1 Matriz de Confusão – Treinamento Satisfeita ou Insatisfeita

Para o conjunto de treinamento dos respondentes que usam Nuvem quando configurado como “com poda”, o algoritmo classificador J48 apresentou os seguintes resultados na matriz de confusão. Essa interpretabilidade pode ser vista na Tabela 2 com o resultado do conjunto de treinamento dos respondentes que usam nuvem.

TABELA 5: Matriz de Confusão – Treinamento Satisfeita ou Insatisfeita

A	B	
42	14	Não
19	17	Sim

Fonte: próprio autor.

Na interpretabilidade da matriz temos: na linha diagonal principal, representada pelos valores 42 (quarenta e dois) e 17 (dezessete), encontra-se o total de 59 (cinquenta e nove) instâncias classificadas corretamente; a mesma operação é realizada para a diagonal transversal, representada pelos valores 19 (dezenove) e 14 (quatorze), somando 33 (trinta e três) instâncias que foram classificadas incorretamente, gerando um total geral de 92 (noventa e dois) instâncias para ambas as diagonais.

- ✓ Instâncias classificadas corretamente com porcentagem  $(59/92)100=64,1304\%$  de exatidão.

- ✓ Instâncias classificadas incorretamente com porcentagem  $(33/92)*100=35,8696\%$  de inexatidão.

Em geral, pode-se concluir que o "erro médio" seria de 35.8696%. Assim, prevê-se que, quando se utilizar o modelo para classificar com poda, erra-se em 35,8696% dos casos.

## 5.2 Resumo do Capítulo

O capítulo de análise é considerado o principal nesta pesquisa. Nele foi possível descrever a análise dos Dados a partir da informação coletada no capítulo anterior (Coleta dos Dados). Para o aprendizado automático, foi aplicada a tarefa de classificação por árvore de decisão, fazendo uso do algoritmo inteligente J48. Procurou-se treinar os conjuntos utilizando a configuração com poda, para se ter uma dimensão da porcentagem de acertos ou erros nos treinamentos aplicados, esses treinamentos foram separados por formulários cada um com um propósito: se usa Nuvem ou não. As distribuições dos ramos das árvores estão dispostas de acordo com o tipo de treinamento realizado. Os resultados foram descritos utilizando a identificação da folha como "R=Resultado", como por exemplo, R1, R2, R3...Rn. Após essa classificação dos resultados, separaram-se os melhores casos e para estes descreveram-se os atributos considerados mais determinantes para as duas linhas de investigação da abordagem da pesquisa. O capítulo seguinte traz as Considerações Finais, dispondo a finalização deste trabalho dissertativo e apresentando as ameaças aos resultados, os trabalhos correlatos, perspectivas de trabalhos futuros, descobertas da pesquisa e conclusões.

## 6 CONSIDERAÇÕES FINAIS

Este capítulo tem como objetivo realizar a finalização da pesquisa abordada em três seções: na seção 6.1, são apresentadas as ameaças dos resultados, identificando quais fatores podem ter interferido nos resultados; na seção 6.2, apresentam-se os trabalhos correlatos, fazendo comparações de pesquisas já realizadas com essa pesquisa, além dos trabalhos futuros, que se apresentam como extensão da atual pesquisa; na seção 6.3, revelam-se as descobertas da pesquisa; e, por fim, a seção 6.4 traz as conclusões.

### 6.1 Ameaças aos resultados

A pesquisa foi desenvolvida com variáveis diversificadas e categóricas através de um formulário *Web*. Diante disso, é possível que alguns fatores possam ter influenciado os resultados, como: obrigatoriedade das respostas, paridade das perguntas, quantidade de instâncias processadas, uma resposta por empresa, escopo da pesquisa, divulgação nas redes sociais e ausência de alternativa como resposta. Cada um desses fatores é apresentado de maneira detalhada a seguir.

**Obrigatoriedade das respostas:** embora tenha sido realizado o Pré-teste, após a liberação da pesquisa propriamente dita, algumas inconformidades foram identificadas na informação coletada, gerando dificuldade na análise dos resultados finais. Identificou-se que o critério de obrigatoriedade das respostas culminou no número de 140 (cento e quarenta) respondentes com respostas incompletas, o que resultou no descarte dessas informações. Além disso, as respostas realizadas em ambos os formulários – em um total de 25 (vinte e cinco) respondentes – também foram desconsideradas.

**Paridade das perguntas:** o modo como foi desenvolvido o questionário não foi projetado para possuir perguntas correspondentes em relação aos participantes que não estão utilizando Nuvem com os participantes que utilizam algum tipo de serviço em Nuvem. A mesma pergunta realizada para quem não usa Nuvem seria a mesma para quem utiliza Nuvem.

**Quantidade de instâncias processadas:** alguns atributos não tiveram ocorrências como resposta, em consequência disso o resultado apresentado foi considerado insignificante. A dissociação desses atributos ocasionou o desprezo por

parte do algoritmo de classificação, o que pode ter influenciado no resultado esperado.

**Uma resposta por empresa:** é possível que a pesquisa possa ter sido respondida por mais de um representante dentro da mesma empresa, o que pode ter provocado interferência no resultado geral da pesquisa.

**Escopo da pesquisa:** o escopo definido com 14 (quatorze) questionamentos pode não ter abrangido alguns questionamentos considerados importantes para a pesquisa, como a existência de profissionais especializados para atuar na área de Computação em Nuvem e portabilidade da Nuvem como benefício que os provedores poderiam oferecer.

**Divulgação nas redes sociais:** a escolha da *Web*, juntamente com as redes sociais, pode não ter sido uma boa escolha, pois respondentes que não são ligados a empresas podem ter respondido os formulários, o que pode ter alterado os resultados.

**Ausência de alternativa como resposta:** para o questionamento específico sobre “segmento de atuação”, observou-se que a alternativa categórica “ATACAREJO”, que é uma nova bandeira de mercado, não foi tratada, o que pode ter influenciado no resultado do aprendizado.

Em geral, conclui-se que, mesmo com essas ameaças aos resultados, as descobertas reveladas foram consideradas satisfatórias, já que, conforme pode ser verificado na seção 6.4, os objetivos da pesquisa foram alcançados.

## 6.2 Trabalhos correlatos e trabalhos futuros

Para maior clareza na exposição dos trabalhos correlatos e futuros, apresentam-se na subseção 6.2.1 os trabalhos correlatos e na subseção 6.2.2, os trabalhos futuros.

### 6.2.1 Trabalhos correlatos

A decisão de migrar para o ambiente em Nuvem é uma atitude que deve ser tomada com muita segurança e consistência. A reflexão e o discernimento são indispensáveis no processo decisório, permitindo a verificação do contexto da empresa e a possibilidade de mudança e, mais ainda, a extensão para o futuro da

organização como um todo. A mudança feita de forma não planejada muito provavelmente acarreta diversos problemas, sobretudo financeiros.

Essa discussão pode justificar os argumentos acima elaborados, sustentados a partir de trabalhos relacionados com a investigação do processo migratório para a plataforma Nuvem no Brasil, interesse deste trabalho dissertativo.

No trabalho dissertativo de Ramalho (2012), realizou-se uma pesquisa sobre a adoção das empresas brasileiras do ambiente em Nuvem, a qual identificou as características dos serviços de Computação em Nuvem utilizados no Brasil e o seu grau de aderência em cada serviço. Durante o processo de investigação, ele fez um levantamento das empresas brasileiras que adotam o ambiente Nuvem e conseguiu mapear as características dos serviços empregados por essas organizações, utilizando para isso alguns questionamentos, tais como os que seguem:

(1) Quais serviços de Computação em Nuvem são oferecidos no mercado e quais as suas características? (2) Quais as motivações das organizações na escolha do fornecedor de Computação em Nuvem? e (3) Qual a importância para a organização dos serviços de Computação em Nuvem contratados? (RAMALHO, 2012, p. 04).

No processo de terceirização de um serviço, devem ser observados os impactos nas organizações e as estratégias de terceirização, tomando como exemplo algumas categorias de empresas, como Suporte, Fábrica, Alinhamento e Estratégia.

- **Suporte:** firmas desta categoria são caracterizadas por terem baixa dependência operacional e estratégica de TI. Para este caso, recomenda-se a terceirização, especialmente para firmas de grande porte.

- **Fábrica:** o perfeito funcionamento das operações de TI é crucial para o desempenho das atividades organizacionais. Para firmas neste quadrante, a recomendação é terceirizar, a menos que sejam firmas de grande porte e com processos gerenciais maduros.

- **Alinhamento:** nesta categoria, as iniciativas de negócio proporcionadas pela TI são críticas para o posicionamento estratégico da organização no futuro. As firmas deste quadrante devem combinar serviços terceirizados com não terceirizados. Na Computação em Nuvem, a estratégia mais usual é terceirizar serviços não estratégicos, tais como serviços de hospedagem e e-mail corporativo, por exemplo.

- **Estratégica:** firmas neste quadrante também devem combinar serviços terceirizados com não terceirizados.

Os dados apresentados por Ramalho (2012) foram coletados por meio de um questionário enviado para os respondentes por e-mail publicado em fórum de TI e conteve 96 casos. A amostragem utilizada é não probabilística, pois os resultados tendem a ser não generalizados. Todavia, como o estudo é exploratório, seus resultados possibilitam a formação de uma base para os futuros trabalhos de natureza explicativa. Comparado com a pesquisa de investigação deste trabalho, pode-se diferenciar o direcionamento da investigação realizada da seguinte forma: em vez de se investigar o impacto dessa terceirização, abordaram-se aqui os motivos da não adesão/satisfação da Nuvem baseado no custo.

Um mapeamento sistemático foi realizado por Carvalho (2012). Nessa pesquisa, o autor procurou realizar uma investigação na área de Computação em Nuvem sintetizando informações importantes. A proposta do autor foi que, a partir do estudo sistemático, com o propósito de apresentar mecanismos que auxiliem os interessados na adesão ao ambiente Nuvem, alcançasse-se um mapeamento sistemático sobre Computação em Nuvem. Foram realizadas buscas automáticas, pelas quais ele conseguiu reunir 2.977 (dois mil novecentos e setenta e sete) títulos, dos quais 301 (trezentos e um) foram identificados como relevantes e classificados de acordo com oito perguntas de pesquisa. Seguindo alguns critérios de relevância e exclusão, os questionamentos realizados foram estes:

QP1 - Que desafios são encontrados em relação a problemas econômicos?

QP2 - Quais problemas e soluções foram encontrados quanto a garantia de serviço?

QP3 - Qual o impacto social de Computação em Nuvem?

QP4 - Quais desafios foram encontrados em relação à concepção de serviços em um ambiente de Computação em Nuvem?

QP5 - Quais principais desafios foram encontrados em relação à propriedade elástica de Computação em Nuvem?

QP6 - Quais os principais problemas e soluções acerca do armazenamento em Computação em Nuvem?

QP7 - Como é feito o monitoramento do uso de recursos em Computação em Nuvem?

QP8 - Quais os principais problemas quanto a segurança em Computação em Nuvem? (CARVALHO, 2012, p. )

Os critérios de inclusão da pesquisa foram:

- O trabalho explorasse a área de Computação em Nuvem como foco principal;
- O trabalho relacionasse Computação utilitária com Computação em Nuvem;
- O estudo comparasse Computação em Nuvem e Computação em Grade.

Os critérios escolhidos para exclusão foram:

- Estudos sem foco na área de Computação em Nuvem;
- Estudos duplicados;
- *Keynotes*, *Whitepapers* e apresentações.

Para Carvalho (2012), os indícios do mapeamento encontrados nesta pesquisa apontam que o interesse direcionado à Computação em Nuvem pela possibilidade de baixo custo inicial é fundamentado em resultados concretos e positivos. No entanto, toda a transparência e comodidade (serviços *on demand*) proporcionada pela Computação em Nuvem mascaram a complexidade do uso de estratégias e tecnologias adotadas para implantar serviços nesse modelo. Esse trabalho é tomado como base para este trabalho devido às referências produzidas a partir do estudo sistêmico do autor, de forma que citam-se fontes confiáveis no trabalho atual.

Um estudo sistemático sobre a elasticidade da Computação em Nuvem, realizado por Emanuel *et al.* (2013), publicado no 31º Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC), em 2013, apresenta a capacidade de adicionar ou remover recursos, sem interrupções e em tempo de execução, para lidar com a variação da carga. De acordo com Garg *et al.* (2012), elasticidade é a capacidade de um serviço escalar, durante períodos de pico, caracterizada pelo tempo médio para expandir ou contrair a capacidade do serviço e capacidade máxima do serviço.

Nessa pesquisa, Emanuel *et al.* (2013) fazem a análise de desempenho que envolve coleta de dados formais e informais para ajudar os clientes e patrocinadores a definir e alcançar seus objetivos. Desse modo, apresentam várias perspectivas sobre um problema ou oportunidade, determinando o direcionamento para barreiras ou desempenho bem-sucedido, e propõem um sistema de solução com base no que é descoberto. Inclui: medição, modelagem, estatística, projeto experimental, simulação, teoria das filas etc. Em geral, são identificados três aspectos: sistema – qualquer conjunto de hardware, software e firmware; métricas – critérios utilizados

para avaliar o desempenho do sistema ou componentes; e cargas de trabalho – requisições realizadas pelos usuários do sistema.

Também no estudo é realizada uma abordagem para avaliação de desempenho, levando em consideração tarefas como: defina metas e o sistema; liste os serviços e seus resultados; selecione métricas; liste parâmetros; selecione fatores para estudo; selecione técnica de avaliação; selecione carga de trabalho; projete os experimentos; analise e interprete os dados; apresente os resultados.

Ao final do processo, foi apresentada uma metodologia para revisão sistemática simplificada, em que foi realizado todo o planejamento da pesquisa e, em seguida, a consolidação dos resultados. Para os autores, os benefícios da revisão foram enormes, como: aspectos relacionados às formas de realizar análise de desempenho, ferramentas, *benchmarks*, cargas de trabalho, métricas e tendências de pesquisa em elasticidade em Computação em Nuvem, com aspectos de análise de desempenho. A abordagem realizada nessa pesquisa segmenta-se na capacidade elástica de adicionar e remover recursos na Computação em Nuvem, a partir da qual se pode ter uma visão do funcionamento dos recursos ofertados pela Nuvem e entender melhor os resultados da pesquisa atual sobre a satisfação das empresas levando em conta esses benefícios.

Jacquet e Cavassana (2012), na linha de pesquisa que discute algumas das tendências relevantes do mercado de Computação em Nuvem, realizam um trabalho com o propósito de identificar pontos específicos que refletem a experiência na utilização da Computação em Nuvem, no universo das pequenas, médias e grandes empresas. Também apresentam um breve histórico do desenvolvimento de tecnologias e realiza uma pesquisa de campo com empresas que já implantaram essa nova ferramenta. Como instrumento de coleta de dados, foi utilizado o questionário, composto de perguntas sobre pontos específicos que refletem as experiências das empresas. Foram consultadas empresas de pequeno, médio e grande porte.

O questionário continha 25 (vinte e cinco) questões de múltipla escolha e questões abertas, focalizando 5 (cinco) aspectos: dados da empresa; motivos da escolha; impactos/custos pós-implementação; vantagens/desvantagens aos usuários finais; considerações.

A pesquisa envolveu 300 (trezentas) empresas que adotaram Computação em Nuvem no período de 2010/2012. De fato, apenas 60 (sessenta) empresas se



mostraram interessadas em responder o questionário e isso foi conseguido por meio de contato telefônico.

Como conclusão mais relevante, destacam-se as diferenças entre empresas brasileiras que ainda sofrem com a falta de conhecimento sobre o tema, enquanto a principal preocupação na Europa e Estados Unidos é a segurança. O trabalho concluiu que o Brasil ainda está na etapa inicial de implantação da Computação em Nuvem. Esta pesquisa permitiu conhecer as variáveis de escolha da tecnologia e os seus impactos no negócio do cliente. Destacam-se quatro pontos, a saber: fundamentação teórica sobre a evolução computacional até Computação em Nuvem; definição das características e camadas de Computação em Nuvem; compreensão das principais variáveis prós e contras desta tecnologia; comparação dos dados obtidos com o referencial teórico analisado.

Jacquet e Cavassana (2012) concluíram que as razões e barreiras para uma empresa adotar a solução de Computação em Nuvem variam conforme o seu tamanho, visão estratégica, disponibilidade de recursos, necessidades, entre outras questões. Entre as principais barreiras destaca-se a adaptação cultural, diagnosticada como o maior empecilho para a implementação da plataforma, motivo que pode estar relacionado à falta de informação e conhecimento sobre Computação em Nuvem, pois se trata de uma tecnologia recente e em constante evolução.

Ao comparar as pesquisas relacionadas, identificou-se que o diferencial da pesquisa realizada nesta dissertação são as linhas de investigações (identificação dos motivos de não usar Nuvem e satisfação de quem a utiliza), assim como a descoberta e extração do conhecimento adquirido através do aprendizado automático.

#### 6.2.2 Trabalhos Futuros

Como proposta de extensão para esta pesquisa, pode-se realizar uma investigação sobre algumas informações pontuais que não foram abordadas nesta pesquisa dissertativa. A seguir são expostas algumas abordagens como proposta para trabalhos futuros.

**A Reaplicação da pesquisa:** o objetivo da reaplicação da pesquisa, além de atualização dos dados, é o aperfeiçoamento mediante a identificação de outras

variáveis para o melhoramento dos resultados. Para isso, deve ser realizada uma reestruturação das perguntas com um critério de paridade viabilizando a interpretação melhor na descoberta e extração do conhecimento.

**Segurança da Computação em Nuvem:** evidenciar as medidas de proteção tomadas por parte dos provedores de serviços da Computação em Nuvem no que tange à segurança da informação. De acordo com a norma ISO/IEC 27002:2005, suas características básicas da segurança da Informação são:

- **Confidencialidade:** propriedade que limita o acesso à informação tão somente às entidades legítimas, ou seja, àquelas autorizadas pelo proprietário da informação.
- **Integridade:** propriedade que garante que a informação manipulada mantenha todas as características originais estabelecidas pelo proprietário da informação, incluindo controle de mudanças e garantia do seu ciclo de vida (nascimento, manutenção e destruição).
- **Disponibilidade:** propriedade que garante que a informação esteja sempre disponível para o uso legítimo, ou seja, por aqueles usuários autorizados pelo proprietário da informação. O objetivo é investigar se essas características estão sendo cumpridas.

**Conversão de novos serviços:** o dinamismo que a Computação em Nuvem possui permite que novos serviços sejam convertidos (Música, Comunicação, Testes, Segurança, Monitoramento) para esse ambiente. A proposta é investigar esses novos serviços, como essas conversões acontecem e como elas funcionam.

### 6.3 Descobertas da pesquisa

Da forma como foi elaborada a pesquisa apresentam-se duas linhas de investigações, logo, geram-se duas hipóteses (uma acerca da contratação ou não da Nuvem e a outra a respeito da satisfação ou insatisfação com a Nuvem). "Uma hipótese é uma suposição que se faz na tentativa de explicar o que se desconhece" (BARBOSA, 2010, p. 75).

Em um contexto geral, a aplicação de aprendizado automático para extração de conhecimento através dessa pesquisa possibilitou prever/descobrir os indícios pelos quais algumas empresas podem estar propensas na contratação ou não dos

serviços de Computação em Nuvem no Brasil. Adicionalmente, possibilitou identificar os motivos, para aquelas que já a usam, sua satisfação ou insatisfação em relação aos serviços utilizados no Brasil.

Cada uma das hipóteses é discutida com mais detalhes nas subseções a seguir.

### 6.3.1 Hipótese de contratar ou não

Inicialmente, as primeiras descobertas realizadas pelo aprendizado referem-se ao treinamento com empresas que não utilizam serviço em nuvem presente na Tabela 3.

TABELA 6: Casos classificados com 100% de acertos

Treinamentos classificados com 100% de acertos							
Seq	Treinamento	Combinação de Atributos			Folha	Instâncias	Classificador
1	Não usa Nuvem	Faturamento: >=2 mi e <=10 mi	Segmento: Varejista	Motivo: Confiabilidade	R15	18	NÃO
2	Não usa Nuvem	Faturamento: >=10 Mi <= 1 Bi	Segmento: Varejista	Motivo: Garantia Entrega	R13	16	SIM
3	Não usa Nuvem	Faturamento: <= 500 mil	Segmento: Outros	Total Colaborador: Entre 21 e 50	R26	14	NÃO
4	Não usa Nuvem	Faturamento: >=10 Mi <= 1 Bi	Segmento: Prestação Serviço	Motivo: Qualidade Serviço	R52	12	SIM

Fonte: próprio autor.

Descobriu-se que, quando combinados os atributos de empresas que faturam entre 2 milhões até 10 milhões atuando no varejo, estas, por sua vez, não estão propensas a contratarem os serviços de nuvem, tendo como motivo a confiabilidade da informação e o faturamento. Isso leva a crer que são empresas de pequeno porte que não despertaram interesse na Computação em Nuvem, mesmo com os benefícios que ela oferece.

Quando combinados os atributos de empresas que possuem faturamento entre 10 milhões e 1 bilhão e atuam como varejistas, descobriu-se que elas estão propensas a contratar os serviços, por encontrarem na garantia de entrega um bom motivo (Sequência 2 da Tabela 3).

Sobre as empresas que faturam menos de 500 mil, de Outros segmentos e que possuem entre 21 (vinte e um) e 50 (cinquenta) colaboradores, descobriu-se que elas estão propensas a não contratar os serviços em Nuvem (Sequência 3 da Tabela 3). Trata-se de empresas de pequeno porte, nas quais a Receita Operacional Bruta (ROB) nesse patamar não possui investimento para contratação da Nuvem. De acordo com Saraceni *et al.* (2011), mais de 90% das empresas brasileiras formalmente registradas são de pequeno e médio porte.

Outra descoberta diz respeito às empresas que faturam entre 10 milhões e 1 bilhão e que são prestadoras de serviços. Identificou-se que elas estão propensas a contratar os serviços da Computação em Nuvem. O treinamento revelou que a qualidade do serviço oferecido pelos provedores é o que as motivam.

No aprendizado em que o treinamento apresenta erros e acertos, tem-se o detalhamento na Tabela 4.

TABELA 4: Casos classificados com Acertos/Erros (Contrataria ou Não Nuvem)

Treinamentos classificados com Acertos e Erros									
Seq	Treinamento	Combinação de Atributos			Folha	Instâncias (Acertos e Erros)		% Erros	Classificador
1	Não usa Nuvem	Faturamento: >=10 Mi <= 1 Bi	Segmento: Prestação Serviço	Motivo: Segurança da Info	R14	8	1	12,5 %	NÃO
2	Não usa Nuvem	Faturamento: < 500 mil	Segmento: Outros	Total Colaborador: < 20 colaboradores	R28	12	2	16,66 %	NÃO

Fonte: próprio autor.

Nos casos em que o aprendizado proporciona Acertos e Erros, foi possível observar que, quando combinadas as variáveis faturamento entre 10 milhões e 1 bilhão e empresas do segmento prestadoras de serviços, não há predisposição à contratação da Nuvem por entenderem que esse ambiente não possui segurança. Embora a margem de erro possua 12,50% das instâncias classificadas incorretamente, é considerado aprendizado bem-sucedido (Sequência 1 da Tabela 4).

Para a combinação das empresas que faturam menos que 500 mil e que são de “Outros” segmentos com menos de 20 colaboradores, também há uma tendência a não contratarem os serviços da Nuvem, uma vez que foram identificados 16,66% das instâncias classificadas incorretamente (Sequência 2 da Tabela 4). Para esse caso acredita-se no faturamento e porte da empresa como fatores determinantes.

### 6.3.2 Hipóteses de Satisfação ou Insatisfação

Para os treinamentos realizados com o propósito de identificar os motivos de satisfação ou insatisfação das empresas que já se encontram utilizando a Computação em Nuvem, as descobertas do aprendizado são esquematizadas nas Tabelas 5 e 6.

TABELA 5: Casos com 100% de acertos (Satisfação ou Insatisfação)

Treinamentos classificados com 100% de acertos								
Seq	Treinamento	Combinação de Atributos				Folha	Instâncias	Classificação
1	Usam Nuvem	Faturamento: >1 Bi	Segmento: Prestação Serviço	Característica: Preço		R10	22	Satisfeita
2	Usam Nuvem	Faturamento: >=10 Mi <= 1 Bi	Seg. Nuvem Publica: Muito confortável	Modelo Serviço: SaaS		R14	9	Satisfeita
3	Usam Nuvem	Faturamento: >=10 Mi <= 1 Bi	Seg. Nuvem Publica: Muito confortável	Modelo Serviço: DaaS		R17	5	Insatisfeita
4	Usam Nuvem	Faturamento: >1 Bi	Segmento: Prestação Serviço	Característica: Segurança	Redução Custo: Menos que 20%	R25	9	Insatisfeita

Fonte: próprio autor.

Assim como a descoberta anterior relacionou a contratação ou não dos serviços da Nuvem, as presenças dos atributos de Faturamento e Segmento também são predominantes nessa classificação. *A priori*, descobriu-se que, quando combinadas as variáveis faturamento acima de 1 bilhão e empresas prestadoras de serviços, elas encontram-se satisfeitas, tendo no preço seu maior motivo (Sequência 1 da Tabela 5).

Uma outra importante descoberta foi a identificação de que, quando combinadas as variáveis faturamento entre 10 milhões e 1 bilhão e pertença ao grupo de empresas que estão utilizando o modelo de Nuvem Pública, essas encontram-se satisfeitas com a utilização do modelo de Software como Serviço-SaaS (Sequência 2 da Tabela 5). Enquanto que empresas que possuem o mesmo faturamento e utilizam modelo de Nuvem Pública estão insatisfeitas (Sequência 3 da Tabela 5), porque o modelo utilizado é o de Banco de dados como Serviço (DaaS). É provável que alguma variável externa tenha influenciado nesse resultado e que o aprendizado não foi capaz de identificá-la.

As combinações das variáveis faturamento acima de 1 bilhão e empresas na área de prestação de serviços demonstraram a existência de insatisfação com a utilização da nuvem, sendo a segurança e redução de custo (menos que 20%) os maiores motivos para essa insatisfação (Sequência 4 da Tabela 5).

O aprendizado em que o treinamento apresenta erros e acertos é demonstrado pela Tabela 6 (seis) a seguir.

TABELA 6: Casos classificados com Acertos/Erros (Satisfação ou Insatisfação)

Treinamentos classificados com Acertos e Erros										
Seq	Treinamento	Combinação de Atributos				Folha	Instâncias (Acertos e Erros)		% Erros	Classificador
1	Usam Nuvem	Faturamento: >1 Bi	Segmento: Prestação Serviço	Características: Disponibilidade		R11	19	1	5.20 %	Insatisfeita
2	Usam Nuvem	Faturamento: > 1 Bi	Segmento: Prestação Serviço	Característica s: Segurança	Redução de custo: Aprox 30%	R23	4	1	25.00 %	Insatisfeita
3	Usam Nuvem	Faturamento: >=10 Mi <= 1 Bi	Seg. Nuvem Publica: Muito confortável	Modelo Serviço: IaaS	Total Colaborador: >1000	R26	30	5	16.66 %	Satisfeita

Fonte: próprio autor.

Revelou-se que quando o treinamento combinou os atributos faturamento acima de 1 bilhão tratava-se de empresas na área de prestação de serviços, estas encontravam-se insatisfeitas em decorrência da disponibilidade dos serviços oferecidos. A presença de 5,20% das instâncias classificadas incorretamente não considera esse treinamento como ruim (Sequência 1 da Tabela 6). Outra descoberta identificada nesse treinamento é que, quando combinados os mesmos atributos da sequência 1 dessa mesma tabela, os motivos para a insatisfação apontam para a segurança e redução de custo, que, embora tenham alcançado 30% de redução, essas empresas estão insatisfeitas. É provável que a segurança seja o fator determinante (Sequência 2 da Tabela 6).

E, por fim, descobriu-se que a combinação dos atributos faturamento entre 10 milhões e 1 bilhão e empresas que utilizam o modelo de infraestrutura como serviço (IaaS) associada ao tipo de Nuvem Pública geram satisfação nas empresas com os serviços, embora o treinamento tenha apresentado 16.66% das instâncias classificadas incorretamente. Pode-se concluir que, pela quantidade de colaboradores associada ao faturamento, tem-se uma visão da dimensão dessas empresas, classificando-as como de grande porte (BNDES, 2010). Presume-se que

estas possuam certa maturidade em tecnologia, daí a satisfação em utilizar o modelo infraestrutura como Serviço (IaaS).

Em linhas gerais, todas essas descobertas só foram possíveis mediante a aplicação do aprendizado automático, pois a partir dele pôde-se evidenciar os indícios para tornar as empresas propensas à contratação ou não, assim como os motivos de satisfação ou insatisfação de estarem utilizando nuvem.

## 6.4 Conclusões

Tem-se escrito muito sobre Computação em Nuvem e sua importância dentro das organizações no que tange à redefinição de um cenário ideal na TIC. A Computação em Nuvem também está redefinindo a forma como as empresas passaram a se relacionar com a informação. É provável que a “dependência” dos profissionais de espaço/estrutura física esteja com seus dias contados, uma vez que a facilidade com que um ambiente em Nuvem pode disponibilizar esses recursos tornou mais práticas, rápidas e flexíveis as atividades das empresas.

Este trabalho realizou um estudo através de um *Survey*, com o propósito de investigar os indícios pelos quais algumas empresas podem estar propensas a contratarem ou não os serviços de Computação em Nuvem no Brasil. Adicionalmente, identificar os motivos, para aquelas que já a usam, de sua satisfação ou insatisfação em relação aos serviços de nuvem contratados no Brasil.

Pode-se evidenciar alguns indícios que tornaram essas empresas propensas a contratar ou não os serviços da Computação em Nuvem. Primeiro, destacam-se as empresas que estão propensas, sendo os indícios a garantia de entrega e qualidade dos serviços ofertados. Em contrapartida, as empresas que não estão propensas a contratar os serviços da Nuvem têm como indícios o baixo faturamento e a quantidade pequena de colaboradores, associados à falta de confiança e insegurança da informação nos serviços ofertados.

Em relação à outra linha de investigação da pesquisa, que é identificar os motivos de satisfação ou insatisfação dos contratantes da Nuvem, podem-se destacar as descobertas dos motivos de satisfação a partir do treinamento em que foi revelado que o preço da Nuvem associado aos modelos de Infraestrutura e Software como Serviço são os indícios dessa satisfação. Por outro lado, descobriu-se que os motivos de insatisfação estão associados à segurança, disponibilidade

dos serviços e redução de custo. A pesquisa revelou que a segurança influencia sobremaneira nos resultados, tanto para tornar as empresas propensas à contratação, quanto nos motivos de insatisfação, o que permite concluir que a segurança não está sendo vista com bons olhos pelas empresas que estão utilizando os serviços da Computação em Nuvem, assim como por aquelas que pretendem utilizar.

Em geral, tanto nas descobertas com 100% de acertos como nas que apresentam erros, percebe-se que esses resultados se mostram de maneira diversificada. Isso indica que uma investigação dessa natureza não tem um atributo único sendo estabelecido como um padrão para cada tipo de treinamento, justamente por se tratar de empresas com realidades diferentes.

Sabe-se da importância do aperfeiçoamento deste estudo, pois se deve proporcionar uma informação precisa e contundente para as comunidades profissional e acadêmica, mas esta já é uma iniciativa no sentido de, utilizando a metodologia aqui disposta, apresentar uma discussão a mais acerca da recepção dos serviços oferecidos pela Computação em Nuvem.



## REFERÊNCIAS

ALBERTO, F, H. **Uma abordagem de acordo de nível de serviço para computação em nuvem**. 2010. Dissertação (Mestrado)-Programa de Pós-Graduação em Ciência da Computação, Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis, 2010.

ALECRIM, E. **Cluster**: conceito e características. Disponível em: <<http://www.infowester.com/cluster.php>>. Acesso em: 16 fev. 2016.

APOSTU, A et al. **Study on advantages and disadvantages of cloud computing: the advantages of telemetry applications in the cloud**. Romania: University Politehnica of Bucharest, 2013.

BANCO NACIONAL DE DESENVOLVIMENTO. **Classificação do porte das empresas**. Disponível em: <[http://www.bndes.gov.br/SiteBNDES/bndes/bndes\\_pt/Institucional/Apoio\\_Financeiro/porte.html](http://www.bndes.gov.br/SiteBNDES/bndes/bndes_pt/Institucional/Apoio_Financeiro/porte.html)>. Acesso em: 30 nov. 2015.

BARBOSA, D. **Manual de pesquisa**: metodologia de estudos e elaboração de monografia. 2 ed. São Paulo: Expressão & Arte, 2010.

BAUN, C. et al. **Cloud computing**: web-based dynamic IT services. Heidelberg: Springer-Verlag, 2011.

BORGES, H. et al. **Computação em nuvem**. Disponível em: <<http://livroaberto.ibict.br/handle/1/861>>. Acesso em: 16 nov. 2015.

BREIMAN, L. Randomizing outputs to increase prediction accuracy. **Technical Report 518**, [S.l.], n. 1, May 1998.

BUBLITZ, E. Catching the cloud: managing risk when utilizing computação em nuvem. **Property, Casualty Risk & Benefits Management**, [S.l.], v. 114, n. 39, p. 12-16, 2010.

CAMILO, C. O.; CARLOS, J. S. **Mineração de dados**: conceitos, tarefas, métodos e ferramentas. Goiás: Universidade Federal de Goiás, 2009. Relatório Técnico-UFG-RT-INF\_001-09.

CARISSIMI, A. Virtualização da teoria a soluções. In: SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES, 2008, [S.l.]. **Minicursos...** [S.l.: s.n.], 2008. p. 173-207.

CARRION, A. P. C. E. **Análise comparativa de Api's para portabilidade em nuvem.** 2014. Trabalho de Conclusão de Curso (Graduação)-Universidade Paranaense, Paranavaí, 2014.

CARVALHO, F. J. S. **Um mapeamento sistemático de estudos em cloud computing.** 2012. Dissertação (Mestrado)-Universidade Federal de Pernambuco, Recife, 2012.

CASTRO, L. C.; BRAGA, A. P. Aprendizado supervisionado com conjuntos de dados desbalanceados. **Sba Controle & Automação**, Campinas, v. 22, n. 5, p. 441-466, set./out. 2011.

CHAPELLE, O.; SCHÖLKOPF, B.; ZIEN, A. (Ed.). **Semi-supervised learning.** Cambridge: MIT Press, 2006.

COLVERO, T. A.; DANTAS, M.; CUNHA, D. P. Ambientes de clusters e grids computacionais: características, facilidades e desafios. In: CONGRESSO SUL CATARINENSE DE COMPUTAÇÃO, 1., 2005, Criciúma. **Anais...** Florianópolis: [s.n.], 2005.

CONTI, F. de. **Grades computacionais para processamento de alto desempenho.** 2009. Trabalho de Conclusão de Curso (Graduação)-Universidade Federal de Santa Maria, Santa Maria, 2009.

CRISTIANO, C. P.; CESAR, E. F. **Metodologia do trabalho científico:** métodos e técnicas da pesquisa e do trabalho acadêmico. Novo Hamburgo: Feevale, 2013.

DE PAULA, M. B. **Introdução automática de árvores de decisão.** 2002. Trabalho de Conclusão de Curso (Graduação)-Universidade Federal de Santa Catarina, Florianópolis, 2002.

EMANUEL, F. C. et al. Elasticidade em computação na nuvem: uma abordagem sistemática. In: SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUÍDOS, 31., 2013. [Trabalhos apresentados]. [S.l.: s.n.], 2013.

ERL, T. **Service oriented architecture**: concepts, technology and design. Crawsford Ville: Prentice Hall, 2005.

FIGUEIREDO, A. B. **IDC**: Brasil. 2013. Disponível em: <<http://br.idclatin.com/prodserv/consulting/>>. Acesso em: 21 mar. 2015.

FONSECA, J. **Introdução de árvore de decisão**: HistClass-proposta de algoritmo não paramétrico. Lisboa: Universidade Nova Lisboa, 1994.

FOSTER, I.; KESSELMAN, C. **The grid 2**: blueprint for a new computing infrastructure. San Francisco: Morgan Kaufmann, 2003.

FREIRE, F. Cezar. **Taurion ameniza as tempestades de questionamentos sobre computação em nuvem**. (Entrevista) 2011. Disponível em: <[http://www.artecom.com.br/revistadigital/downloads/12/link\\_12\\_4047.pdf](http://www.artecom.com.br/revistadigital/downloads/12/link_12_4047.pdf)>. Acesso em: 15 jun. 2015.

GARG, S. K; VERSTEEG, S; BUYYA, R. A framework for ranking of cloud computing services. **Future Generation Computer Systems**, [S.l.], v. 29, p. 1012-1023, 2012.

HARAUZ, J; KAUFMAN, L. M; POTTER, B. Data security in the world of cloud computing. **IEEE Journal on Cloud Computig Security**, [S.l.], v. 7, n. 4, p. 61-64, July/Aug. 2009.

JACQUET, A.; CAVASSANA, H. **Computação em nuvem**: pesquisa de aceitação e implantação no mercado corporativo brasileiro. 2012. Disponível em: <<http://www.portalamericas.edu.br/revista/pdf/ed10/art04.pdf>>. Acesso em: 21 jun. 2015.

KAUARK, F. S; CASTRO, F. M; MEDEIROS, C. H. **Metodologia da pesquisa**: um guia prático. 1 ed. Itabuna: Via Litterarum, 2010.

KUROSE, J. F. **Redes de computadores e a internet**. 5. ed. São Paulo: Pearson, 2010.

LINTHICUM, D. S. **Cloud computing and SOA convergence in your enterprise**: a step-by-step guide. 1. ed. São Paulo: Pearson Education, 2010.

LINTHICUM, D. S. **Sim, a nuvem está substituindo hardware e software corporativos**. 2013. Disponível em: <<http://cio.com.br/tecnologia/2013/12/03/sim-a-nuvem-esta-substituindo-hardware-e-software-corporativos/>>. Acesso em: 07 jan. 2016.

LOUREIRO, S. C. **Segurança da informação**: preservação das informações estratégicas com foco em sua segurança. Brasília: Universidade de Brasília, 2008.

LUGER, G. F. **Inteligência artificial**: estruturas e estratégias para a solução de problemas complexos. 4. ed. Porto Alegre: Bookman, 2007.

MAHMOOD, Z.; HILL, R. **Computação em nuvem for enterprise architectures**. Londres: Springer-Verlag, 2011.

MARI, A. **Cloud adoption grows in Brazil as organizations prepare for tighter budgets**. Capgemini. Disponível em: <<http://www.zdnet.com/article/cloud-adoptiongrows-in-brazil-as-organizations-prepare-for-tighter-budgets/>>. Acesso em: 08 jan. 2016.

MELL, P; GRANCE, T. The NIST definition of cloud computing. **Special Publication 800-145**, 2009. Disponível em: <[http://www.nist.gov/customcf/get\\_pdf.cfm?pub\\_id=909616](http://www.nist.gov/customcf/get_pdf.cfm?pub_id=909616)>. Acesso em: 12 dez. 2015.

MOHAMED, A. **A history of cloud computing**. 2009. Disponível em: <<http://www.computerweekly.com>>.

PEREIRA, A.; PATRÍCIO, T. **SPSS**: guia prático de utilização-análise de dados para ciências sociais e psicologia. 8 ed. Lisboa: Sílabo, 2013.

RAMALHO, C. N. L. R. **Um estudo sobre adoção da computação em nuvem no Brasil**. São Paulo: Universidade de São Paulo, 2012.

REESE, G. **Cloud application architectures**. 1. ed. Sebastopol: O'Reilly Media, 2009.

RUSSELL, S.; NORVIG, P. **Artificial intelligence**: a modern approach. 2.ed. [S.l.]: Prentice Hall, 2003.

SARACENI, A. V.; ANDRADE, J. P. P.; MATOS, S. N. **Computação em nuvem como ferramenta de gestão para pequenas e médias empresas brasileiras**.

2011. Disponível em: <<http://pg.utfpr.edu.br/expout/2011/artigos/15.pdf>>. Acesso em: 15 dez. 2015.

SILVA, N. A. E. **Uma abordagem dirigida por modelos para portabilidade entre plataforma de computação em nuvem**. São Carlos: Universidade Federal de São Carlos, 2013.

SKYINONE. **Levando seu software para a nuvem**. [S.l.]: Ebook, 2014.

SOLANGE, H. L. B., **Análise da completude dos relatos de experimentos em elasticidade na computação em nuvem: um mapeamento sistemático**. Recife: Universidade Federal de Pernambuco, 2015.

STANGE, R. L.; J. NETO, J. Aprendizagem incremental usando tabelas de decisão adaptativas. In: WORKSHOP DE TECNOLOGIA ADAPTATIVA, 5., 2011. [**Trabalhos apresentados**]. [S.l.: s.n.], 2011.

TANENBAUM, A. S. **Rede de computadores**. 5. ed. [S.l.]: Hardcover, 2010.

TAURION, C. **Computação em nuvem: transformando o mundo da tecnologia da informação**. Rio de Janeiro: Brasport, 2009.

\_\_\_\_\_. **É hora de explicar os tsunamis transformadores da TI para os CEOs**. 2015. Disponível em: <<http://cio.com.br/gestao/2015/02/04/e-hora-de-explicar-ostsunamis-transformadores-da-ti-para-os-ceos/>>. Acesso em: 14 jun. 2015.

TELECO. **Market share das operadoras de banda larga fixa no Brasil**. Disponível em: <<http://www.teleco.com.br/blarga.asp>>. Acesso em: 02 mar. 2016.

VECCHIOLA, C.; CHU, X.; BUYYA, R. **Aneka: a software platform for .net-based cloud computing**. 2009. CoRR, abs/0907.4622

VELTE, A. T.; VELTE, J.; ELSENPETER, R. **Cloud computing: computação em nuvem-uma abordagem prática**. Rio de Janeiro: Alta Books, 2010.

VERDERAMI, B. M. Avaliando o uso da computação em nuvem na TI para pequenas e médias empresas brasileiras. **Revista Científica**, [S.l.], v.2, n.1, 2013.

ZUBEN, F. J. V; ATTUX, R. F. **Árvore de decisão**. Campinas: Editora da UNICAMP, 2010.

## APÊNDICES

## APÊNDICE A1 – Formulário da Pesquisa



### Pesquisa sobre adesão da Computação em Nuvem!

"A nuvem é como um grande conjunto de recursos virtualizados de fácil utilização e acessível (como plataformas de hardware, desenvolvimento e / ou serviços). Esses recursos podem ser dinamicamente reconfigurado para se ajustar a uma carga variável (escala), permitindo também uma melhor utilização dos recursos. Este conjunto de recursos é tipicamente explorado por um modelo pay-per-use (pagar para usar) em que as garantias são oferecidas pelo Provedor de Infraestrutura por meios de SLAs personalizados."(EXin-Cloud,2013)

**Vale ressaltar que esta pesquisa possui um caráter científico, a mesma não será utilizada em outros meios senão o acadêmico. Informamos também que o resultado dessa pesquisa será apresentado em seu formato consolidado, e não individual.**

Para esta pesquisa teremos três tipos de formulários:

- I- Dados preliminares (Preenchimento obrigatório).
- II- Empresas que **NÃO** utilizam serviços em Nuvem. (Somente empresas que **NÃO** utilizam deverão preencher)
- III- e Empresas que **UTILIZAM** serviços em Nuvem. (Somente empresas que **UTILIZAM** deverão preencher)


### Seja bem vindo a nossa pesquisa sobre Computação em Nuvem.

Próximo ▶

Sair e apagar o questionário



## APÊNDICE A2 - Formulário das empresas que utilizam Nuvem



**PESQUISA SOBRE COMPUTAÇÃO EM NUVEM**

*Esta pesquisa visa o levantamento de informações para auxiliar no processo de investigação do uso da COMPUTAÇÃO EM NUVEM. Nosso objetivo é levantar informações que nos proporcione uma visão aproximada dos motivos pelos quais as empresas não estão migrando para essa nova plataforma*

0%

100%

**EMPRESAS QUE UTILIZAM COMPUTAÇÃO EM NUVEM**

Estas perguntas são destinadas as empresas/Organizações que utilizam nuvem na atualidade. Para as empresas que utilizam algum tipo de serviço a resposta desse formulário é essencial.

**Qual serviço sua Empresa/Organização utilizado?**  
Escolha a(s) que mais se adequa(m)

☐ Software como Serviço - SAAS

☐ Infraestrutura como Serviço-IAAS

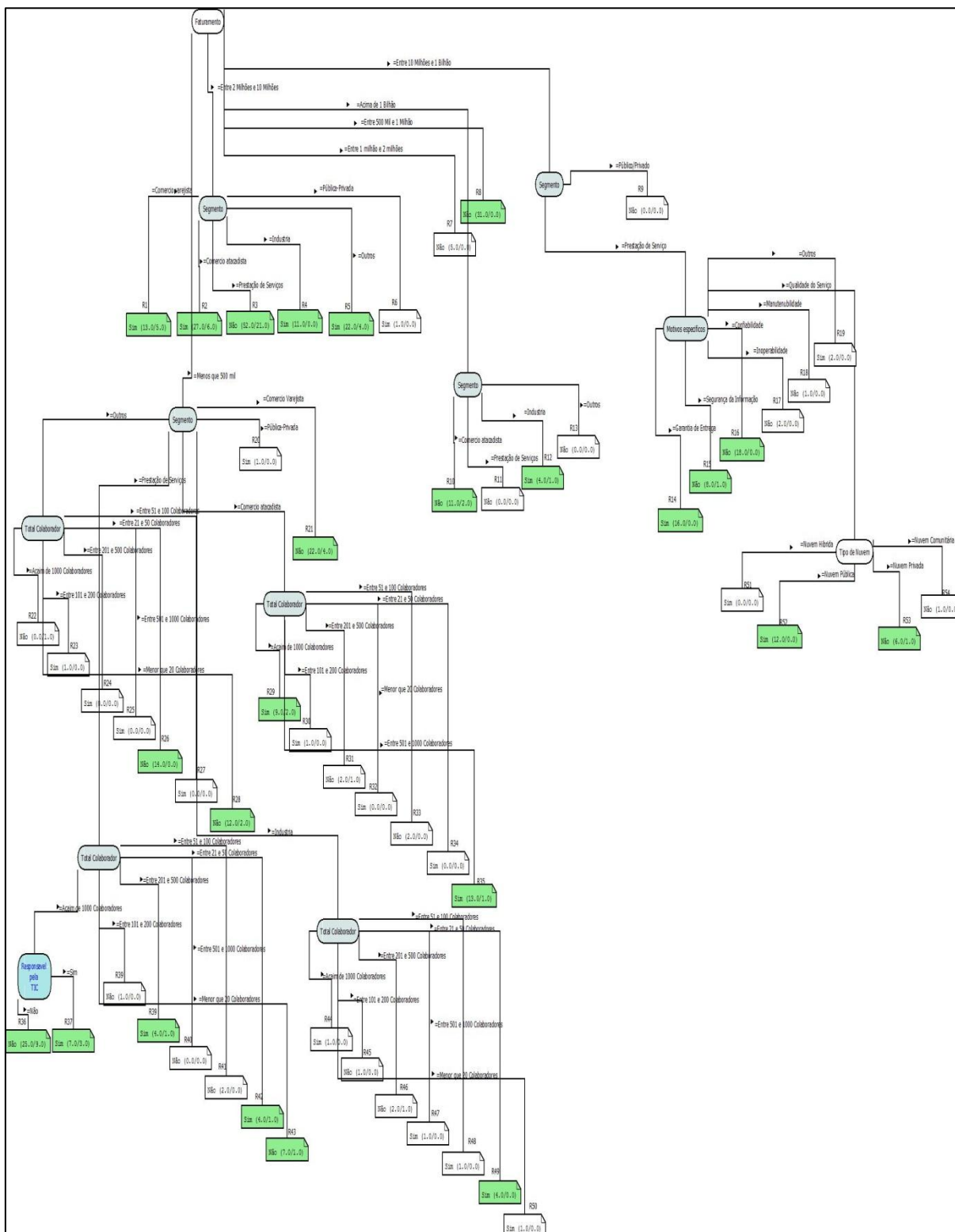
☐ Plataforma como Serviço-PAAS

☐ Armazenamento como Serviço-STAAS

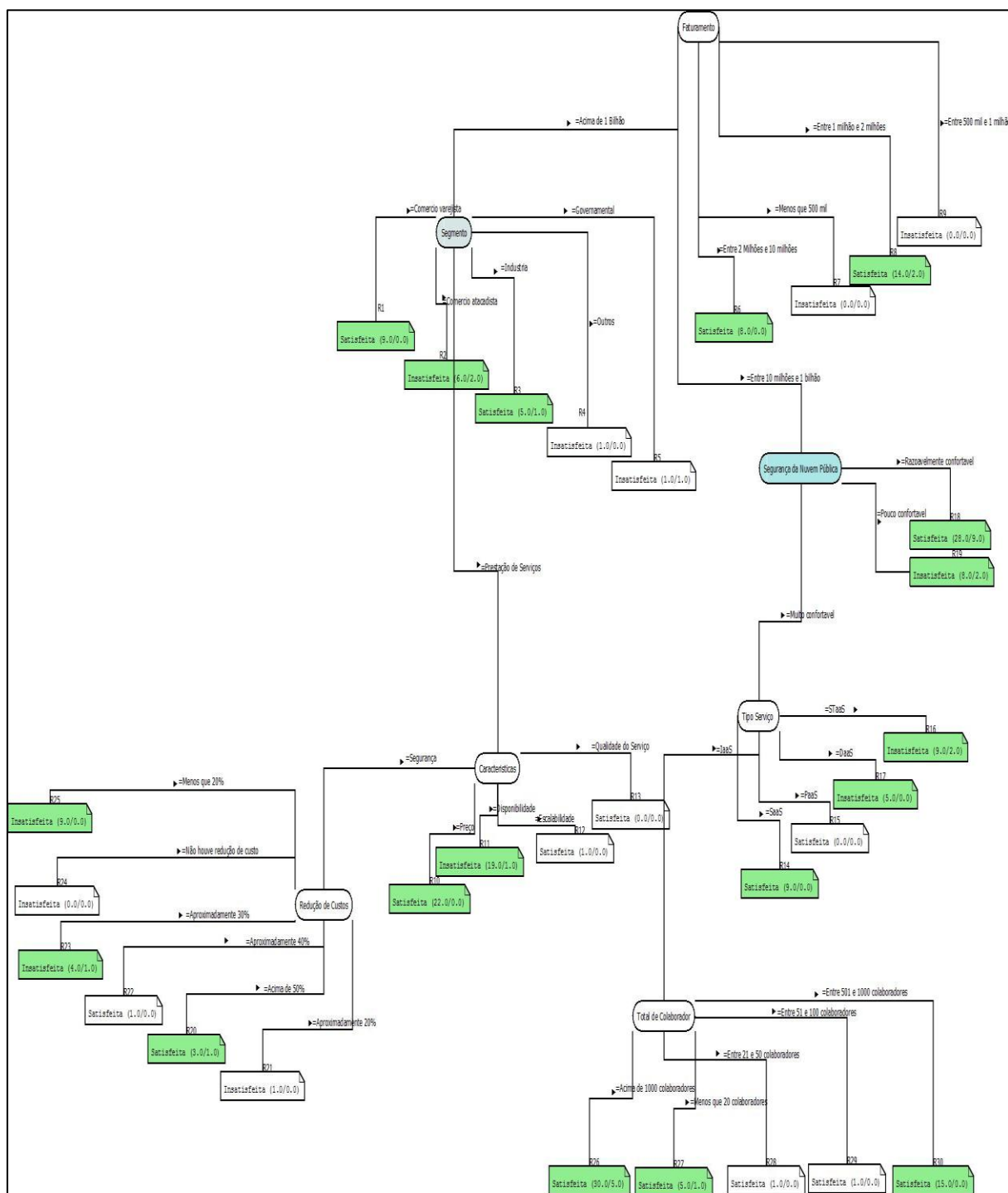
☐ FORENSIC como Serviço-FAAS

☐ Banco de dados como Serviço-DAAS

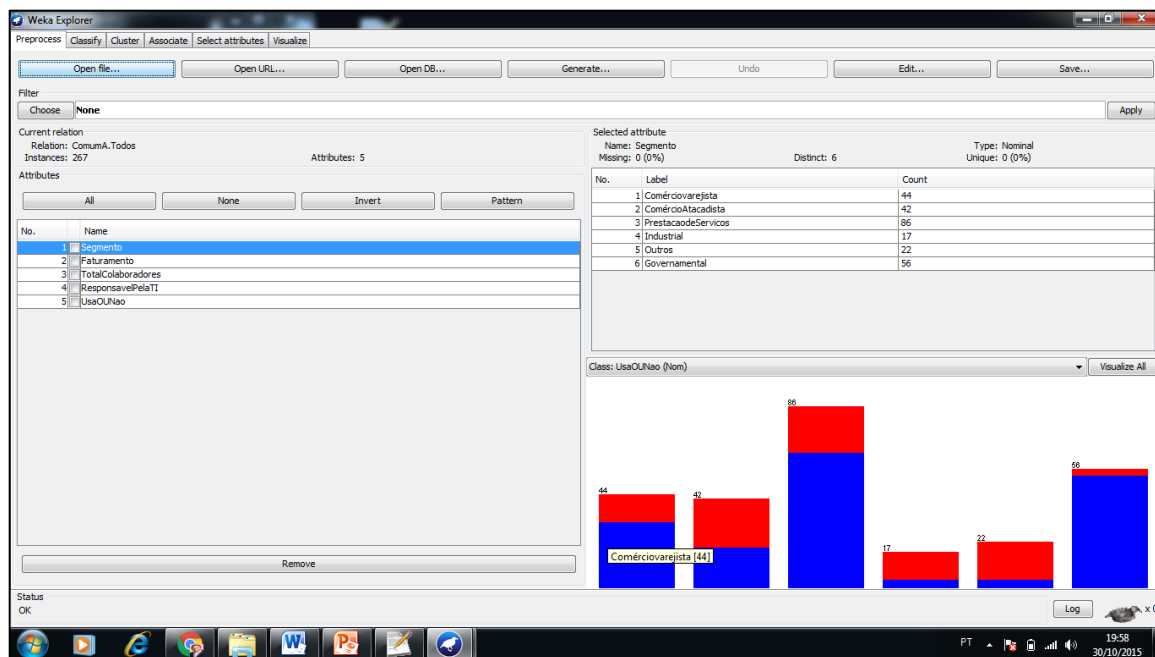
☐ ? Deverá responder qual o tipo de serviço utilizado no momento pela sua Empresa/Organização.



## APÊNDICE B2– Visão Geral da Árvore de decisão – Satisfeita ou Insatisfeita



## APÊNDICE B3 – Treinamento Dados preliminares.



## APÊNDICE B4 – Arquivo Não usa nuvem

```

1 @relation dados.preliminares
2
3 % Dados preliminares
4 @attribute segmento {PrestacaodeServicos, Outros, Industrial,Governamental,Comercio varejista,ComercioAtacadista}
5 @attribute faturamento {Menosque500Mil, Entre500millemilhao, Entre2Milhoese10Milhoes,Entre10Milhoese1Bilhao,Entre1milhaoe2Milhoes,Acimade1Bilhao}
6 @attribute QuantidadeColaborador {Menorque20colaboradores, Entre2leas50colaboradores, Entre5le100colaboradores,Entre50le1000colaboradores,Entre20le500colabo
7 @attribute ResponsavelTI {TRUE, FALSE}
8 @attribute UsaOUNaoUsaNuvm {YES, NO}
9
10 @data
11 Comerciovarejista,Entrelmilhaoe2Milhoes,Acimade1000colaboradores,FALSE,NO
12 ComercioAtacadista,Entre10Milhoese1Bilhao,Acimade1000colaboradores,FALSE,NO
13 ComercioAtacadista,Entre10Milhoese1Bilhao,Acimade1000colaboradores,FALSE,NO
14 PrestacaodeServicos,Entre10Milhoese1Bilhao,Acimade1000colaboradores,FALSE,NO
15 PrestacaodeServicos,Entre10Milhoese1Bilhao,Acimade1000colaboradores,FALSE,NO
16 PrestacaodeServicos,Entre10Milhoese1Bilhao,Acimade1000colaboradores,FALSE,NO
17 PrestacaodeServicos,Entre10Milhoese1Bilhao,Acimade1000colaboradores,FALSE,NO
18 PrestacaodeServicos,Entre10Milhoese1Bilhao,Acimade1000colaboradores,FALSE,NO
19 PrestacaodeServicos,Entre10Milhoese1Bilhao,Acimade1000colaboradores,FALSE,NO
20 PrestacaodeServicos,Entre10Milhoese1Bilhao,Acimade1000colaboradores,FALSE,NO
21 PrestacaodeServicos,Entre10Milhoese1Bilhao,Acimade1000colaboradores,FALSE,NO
22 PrestacaodeServicos,Entre10Milhoese1Bilhao,Acimade1000colaboradores,FALSE,NO
23 PrestacaodeServicos,Entre10Milhoese1Bilhao,Acimade1000colaboradores,FALSE,NO
24 PrestacaodeServicos,Entre10Milhoese1Bilhao,Acimade1000colaboradores,FALSE,NO
25 PrestacaodeServicos,Entre10Milhoese1Bilhao,Acimade1000colaboradores,FALSE,NO
26 PrestacaodeServicos,Entre10Milhoese1Bilhao,Acimade1000colaboradores,FALSE,NO

```

**APÊNDICE B5 – Classificação das Empresas**

<b>Classificação</b>	<b>Receita Operacional Bruta anual (ROB)</b>
Microempresa	Menor ou igual a R\$ 2,4 milhões
Pequena empresa	Maior que R\$ 2,4 milhões e menor ou igual a R\$ 16 milhões
Média empresa	Maior que R\$ 16 milhões e menor ou igual a R\$ 90 milhões
Média-grande empresa	Maior que R\$ 90 milhões e menor ou igual a R\$ 300 milhões
Grande empresa	Maior que R\$ 300 milhões

Fonte: BNDES, 2010 (adaptado).