

EXTRAÇÃO DE CONHECIMENTO EM UM *SURVEY* SOBRE O USO DE COMPUTAÇÃO EM NUVEM NO BRASIL

Hilson Barbosa da Silva, Vinicius Cardoso Garcia

Universidade Federal de Pernambuco
{hbs, vcg}@cin.ufpe.br

Thiago Carvalho de Sousa

Universidade Estadual do Piauí
thiago@uespi.br

Ricardo Andrade Lira Rabêlo, Jefferson Henrique Camelo Soares

Universidade Federal do Piauí
ricardoalr@ufpi.edu.br, jeffersonpi@hotmail.com

Lucas Araújo Lopes

Infoway E-Health Company
lucaslopes@infoway-pi.com.br

Resumo A Computação em Nuvem apresenta um novo conceito de terceirização na contratação de serviços, sendo vista como uma nova possibilidade para a otimização nos volumes dos investimentos em Tecnologia de Informação e Comunicação (TIC). No entanto, muitas empresas brasileiras ainda estão receosas em adotar tal modelo e, para aquelas que já o utilizam, não se conhece o seu grau de satisfação. Nesse contexto, este trabalho aplicou uma pesquisa (*Survey*) que teve como objetivos investigar os indícios pelos quais algumas empresas podem estar propensas a contratar ou não os serviços de Computação em Nuvem, bem como identificar os motivos, para aquelas que já a usam, de sua satisfação ou não em relação aos serviços contratados no Brasil. Após a coleta de dados dos 432 participantes da *Survey*, foi executada uma extração de conhecimentos a partir da aplicação de um aprendizado supervisionado. Como resultado desta análise, descobriu-se importantes informações sobre tal mercado no Brasil, que podem ser utilizadas para subsidiar tomadas de decisão por parte das empresas que ofertam serviços de Computação em Nuvem.

Palavras-chaves: Computação em Nuvem; Extração do conhecimento em Banco de Dados; Aprendizado de Máquina; *Survey*.

Abstract Cloud Computing introduces a new concept of services outsourcing, bringing a new possibility to optimize the investments in Information and Communication Technology (ICT). However, many Brazilian companies are still afraid to adopt this model, and for those who already use it, they do not know their degree of satisfaction. In this context, this study has applied a Survey which aimed to investigate the evidence about why some companies may be prone to subscribe or not the services provided by Cloud Computing, as well as identifying the reasons of the satisfaction or not about the services contracted in Brazil for those that already use them. After collecting data from 432 participants of the Survey, a knowledge discovery was carried from running a supervised machine learning algorithm. As the result of this analysis, it was found important information about this market in Brazil, which can be used to support investment decision by companies that offer Cloud Computing services.

Keywords: Cloud Computing; Knowledge-Discovery in Databases; Machine Learning; *Survey*.

1 Introdução

De acordo com [1], a tecnologia de Computação em Nuvem pode ser considerada, atualmente, como uma tecnologia que introduz novas ideias de negócios nas organizações. Para [2], a Computação em Nuvem é um movimento que está promovendo uma ruptura nos departamentos empresariais de TIC. Ainda para [3], 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. Já para [4], a tecnologia em Nuvem está se disseminando e transformando as empresas e até mesmo aquelas corporações, que se acham distantes dessas transformações, serão afetadas, mais cedo ou mais tarde. De acordo com [5], 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. Entretanto, mesmo sabendo dos benefícios desse tipo de investimento, presume-se que muitas empresas estão temerosas na contratação de serviços e/ou infraestruturas de TIC da Computação em Nuvem. Uma pesquisa realizada em 2014 pela Tech Supply¹ revelou que apesar de todas as facilidades proporcionadas, 43% das empresas brasileiras não se sentem seguras para migrar os seus sistemas para a Computação em Nuvem. Diante disto, pode-se questionar por qual motivo essas empresas não estão contratando os serviços da Computação em Nuvem? E, aquelas que já a adotaram, já têm condições de afirmar se estão satisfeitas ou não, e por quais motivos?

Diante deste contexto geral este artigo apresenta uma investigação dos indícios pelos quais muitas empresas não estão propensas a contratar os serviços de Computação em Nuvem no Brasil, bem como o grau de satisfação em relação aos serviços utilizados daquelas que já a usam. Para tanto, definiu-se uma metodologia de pesquisa exploratória de natureza descritiva e explicativa, com ênfase na abordagem quantitativa, optando-se pela aplicação de um *Survey* online utilizando a ferramenta *LimeSurvey*². Os dados coletados de 432 respondentes foram analisados tendo como suporte a ferramenta WEKA [6], que foi utilizada para extração de conhecimento via aprendizado automático. Esta ferramenta foi configurada para aplicar a técnica de classificação por árvore de decisão com poda utilizando o classificador J48 [7]. Os resultados descobertos proporcionam conhecer com maiores detalhes os motivos pelos quais as empresas brasileiras não aderem a Computação em Nuvem, como por exemplo o baixo faturamento, bem como ter um melhor entendimento sobre qual a razão principal de satisfação (preço reduzido) e de insatisfação (segurança da informação) das empresas que já a adotam.

O restante do artigo está organizado da seguinte forma: na Seção 2 será apresentado o embasamento teórico, com explanações sobre os conceitos de Computação em Nuvem e Aprendizado Automático; na Seção 3 relata-se a metodologia usada, com foco no *Survey* aplicado, enquanto que na Seção 4 se analisam os resultados providos pela extração de conhecimento via aprendizado automático; e por fim a última Seção 5 apresenta as conclusões, bem como trabalhos relacionados e futuros.

2 Referencial Teórico

2.1 Computação em Nuvem

De acordo com [8], a “A fusão dos computadores e das comunicações promoveu uma profunda influência na forma como os sistemas computacionais são organizados”. O novo paradigma de tecnologia, chamado Computação, representa muito bem esse novo cenário gerado. Alguns autores, como [9], definem *Cloud* como sendo uma plataforma de fornecimento de serviços, infraestrutura, plataforma e software sob demanda. Para a AWS³, “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 SLA⁴s personalizados.

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 foi adotada nesta pesquisa:

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 [10].

Para o surgimento da Computação em Nuvem é importante destacar a contribuição de algumas tecnologias precursoras que possibilitaram o seu aparecimento, tais como a Computação Utilitária, Clusterização, Computação em Grade, e Virtualização.

Computação utilitária: de acordo com [11], os conceitos iniciais da Computação em Nuvem 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 na década de 60. No entanto, desde a virada do milênio, o conceito foi reformulado com o surgimento de novas tecnologias no campo da virtualização. Nos tempos atuais, as empresas não têm mais a necessidade de adquirir computadores robustos (servidores) tal como era feito há pouco tempo. Hoje o modelo de Computação Utilitária

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

² Disponível em <http://www.limesurvey.org>

³ Empresa americana provedora de serviços em nuvem. Amazon Web Services. Disponível em: <http://www.aws.amazon.com>. Acesso em : 16 de agosto de 2015.

⁴ Um Acordo de Nível de Serviço (SLA) é uma obrigação contratual entre contratante e seu fornecedor [17].

oferece uma série de benefícios para os prestadores de serviços e usuários [12].

Clusterização: de acordo com [13], “A ideia inicial deste sistema surgiu em 1960, quando a IBM, interligou os seus mainframes e obteve ganho no processamento através do paralelismo”. A Clusterização possui aspectos que reúne dois ou mais computadores para que estes 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. Segundo [14], “Para reforçar estes aspectos, pode-se até mesmo, adotar sistemas operacionais preparados especialmente para *clustering*.”. De acordo com [15], *Cluster* é um tipo de sistema para processamento paralelo e distribuído que consiste de uma coleção de computadores interconectados e trabalhando juntos como um único recurso computacional integrado. Para [16], as soluções desenvolvidas em arquitetura *cluster computing* são usadas 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. Um ambiente clusterizado, geralmente, é desenvolvido para proporcionar recursos computacionais com alto poder de processamento, performance e desempenho.

Computação em Grade: a tecnologia da Computação em Grade possui características semelhantes à Clusterização. De acordo com [13], “Esse sistema computacional surgiu nos anos 90, para suprir com a demanda por alto desempenho computacional”. O termo Computação em Grade é inspirado no conceito de *Power Grid*, que foi idealizado por Ivan Foster [18] no ano de 1998. Esse termo designa o sistema de geração e distribuição de energia. Os participantes dessa arquitetura conectam-se, através de sistemas distribuídos com um gerenciador dos recursos de hardware, rede, conexão com a banda larga e processamento dos dados. A *Grid* possui algumas características semelhantes 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, percebe-se que algumas características se apresentam iguais.

Virtualização: Há dez anos, quando o tema “Nuvem” ainda estava sendo apresentado, seu foco era na oferta de 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, começando com a virtualização de hardware e seus sistemas operacionais, e chegando à 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 de melhorar a utilização dos recursos de servidores *mainframes* (máquinas e servidores de grande porte). Segundo [1], o ambiente virtual possui como característica principal a computação distribuída, virtualizando máquinas, tecnologias e métodos para dar suporte a uma estrutura de softwares. Com a evolução da tecnologia surgiram novos termos como Computação em Nuvem, que faz uso da virtualização.

O NIST [10] apresentou em 2009 um documento que é referência para examinar as características essenciais que a Computação em Nuvem possui. São elas:

- **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;
- **Amplio 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. Para [19], entende-se 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 [20].

Em relação à forma como a Computação em Nuvem é disponibilizada ao usuário final, [10] 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.

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, de acordo com [21], tem-se a seguinte classificação:

- **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;
- **Armazenamento como Serviço (StaaS):** permite o armazenamento de informações como forma de serviços;
- **Banco de dados como Serviço – DaaS:** permite o funcionamento do Sistema de Gerenciamento de Banco de Dados (SGDB) a partir do ambiente em nuvem.

Nessa fundamentação teórica também se aborda a área de conhecimento do aprendizado automático, explanando alguns conceitos de métodos de classificação e aprendizado na seção a seguir.

2.2 Aprendizado Automático

Para [22], a Inteligência Artificial é uma das ciências mais recentes, atualmente abrangendo uma variedade enorme de subcampos, que vão desde áreas de uso geral, como aprendizado e percepção, até tarefas mais específicas, como jogos de xadrez. Aprendizagem de Máquina (*Machine Learning*) é uma técnica que permite que o computador aprenda ou aperfeiçoe seu desempenho em certas tarefas através de algoritmos inteligentes, sendo uma subárea da Inteligência Artificial. Dentro desta subárea, o método de classificação por Árvore de Decisão é um dos mais usados nos treinamentos Supervisionados e Não-supervisionados. Para [23] o aprendizado supervisionado pode ser aplicado tanto para exemplos como contraexemplos, seja o rótulo correto para classificação ou o valor numérico para regressão; isso permite que o algoritmo possa avaliar e validar quão próximo o conhecimento que foi aprendido está do conhecimento correto. Já o treinamento Não supervisionado, não há uma definição de contraexemplos para os dados disponíveis, o que implica na incapacidade, *a priori*, de separar os exemplos a partir de um rótulo bem definido. Sua aplicação tem como objetivo tentar prever valores de variáveis do tipo categóricas que assumem valores finitos e que podem ser ordenadas. O trabalho de [24] faz um detalhamento do 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 [24].

A árvore de decisão também pode ser gerada utilizando a configuração com podagem, utilizando o método por indução e não incremental, segundo [25]. 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.

Ainda para [25], 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. Na classificação por árvore de decisão utiliza-se o aprendizado simbólico, que é detalhado na seção seguinte.

2.2.1 Aprendizado Simbólico

O aprendizado simbólico implica nas adaptações e mudanças do sistema no sentido de permitir a realização das mesmas tarefas nas mesmas circunstâncias de um modo mais eficiente e efetivo para o treinamento. Para a Inteligência Artificial Simbólica, no decorrer da sua evolução, foi adquirindo mecanismos que facilitassem a descoberta automática de conhecimentos, a linguagem de programação Prolog, juntamente, com os sistemas especialistas. Segundo [25] alguns dos métodos de aprendizado simbólico são:

- Aprendizado por analogia/por instâncias. Exemplo: Sistemas baseados em casos - CBR (*Case Based Reasoning*) [7];
- Aprendizado por Indução. Exemplos: Árvores de Decisão - ID3, C4.5, CN2 (IDT – *Induction of Decision Trees*) [27], e ILP - *Inductive Logic Programming (Prolog)* [28].
- Aprendizado por evolução/seleção. Exemplo: Algoritmos Genéticos - GA e GP (*Genetic Algorithms / Genetic Programming*) [29];
- Outros tipos de aprendizado: por reforço (*reinforcement learning*), não supervisionado, bayesiano e por explicações (*explanation based*) [7, 28].

Assim como outras áreas de conhecimento a I.A também possui algumas limitações em seus aprendizados. Para [26], essas limitações são:

- Usualmente assumem que os conhecimentos a serem adquiridos/manipulados e as informações disponíveis são corretos e devem estar completos em relação ao problema (o que dificilmente ocorre);
- São orientadas para manipular informações simbólicas (informações qualitativas), onde os valores numéricos e contínuos (informações quantitativas) representam um problema difícil de ser tratado [26].

3 Metodologia

Para este estudo definiu-se o tipo de pesquisa a ser realizada como exploratória de natureza descritiva e explicativa, com ênfase na abordagem quantitativa, o que possibilita o uso de dados estatísticos para a interpretação dos resultados coletados, bem como para extração automática de conhecimento. Para o procedimento técnico aplicou-se um levantamento por meio de um *Survey*, utilizando-se o instrumento de um questionário online estruturado e fechado com 14 (quatorze) itens, definido da seguinte forma:

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

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.

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.

A pesquisa foi direcionada para alguns segmentos, sendo encaminhado convite de participação da pesquisa para os setores de TIC (Associações, Indústria, Comércio e Tecnologia), Comércio (Varejistas e Atacadistas) e Instituições de Ensino Superior – IES (públicas e privadas). 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.

Para viabilizar as respostas e coleta dos dados, 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. O Gráfico 1, detalha melhor os dados que foram coletados.



Gráfico 1: Resultado Geral

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% responderam o formulário correspondente aos que não utilizam Computação em Nuvem, 92 (noventa e dois) - 21% respondentes afirmaram utilizar algum tipo de tecnologia em Nuvem.

Identificou-se que 25 (vinte e cinco) - 6% responderam a ambos os formulários e 140 (cento e quarenta) - 32% responderam de forma incompleta. A esses dois últimos grupos não será aplicada a análise do material coletado, partindo do princípio que não é possível identificar a resposta correta. Por meio dessa coleta, pode-se ter uma visão aproximada dos indícios que podem tornar essas instituições propensas ou não a contratar a Nuvem, assim como, a identificação de alguns motivos de satisfação ou insatisfação entre as instituições que já utilizam a Computação em Nuvem. Os resultados foram distribuídos de acordo com as perspectivas definidas na metodologia da pesquisa.

4 Análise dos Dados

Para o uso de aprendizado automático faz necessário alguns esclarecimentos em relação aos métodos de aprendizagem que foram aplicados. Para o aprendizado utilizou-se a tarefa de classificação por árvore de decisão com algoritmo de classificação j48 [30], conforme proposto por [31]. Em relação ao método de aprendizagem, utilizou-se aprendizado por indução que permite aprender a partir de conjunto de exemplos. [32]. Após definir o método é fundamental definir o modo de treinamento, o qual foi escolhido o não incremental, que não permite acrescentar um novo exemplo a partir do conjunto de treinamento atual [33]. Na hierarquia do aprendizado utilizou-se o aprendizado supervisionado que possui a figura representativa do “Professor” como medida de classificação dos dados [34]. Em relação ao paradigma de aprendizado aplicou-se o simbólico, que possui no aprendizado um conjunto de representação simbólica [35]. Quanto à configuração dos treinamentos aplicou-se a “com poda”, que consiste em classificar os casos mais significativos do treinamento [36]. E, por fim, definiu-se as variáveis classificadoras de acordo com a abordagem da pesquisa: para a primeira linha de pesquisa, “SIM” que contratariam serviços de Computação em Nuvem ou que “NÃO” contratariam; para a segunda linha de investigação, o classificador “SATISFEITO” com a utilização da dos serviços de Computação em Nuvem ou “INSATISFEITOS”.

Para a etapa de análise, utilizou-se a ferramenta *open source* WEKA 3.6 [6], que foi crucial nesse processo de aprendizado e construção do conhecimento. A princípio, 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 as descobertas através do aprendizado, utilizando-se para isso algoritmos de classificação por árvore de decisão. O primeiro conjunto de treinamento foi projetado no intuito de identificar os participantes que não utilizam a Computação em Nuvem, com o classificador J48 gerando uma árvore para cada configuração utilizada (com poda). Esse processamento gerou uma árvore de 12 (doze) níveis e com 54 (cinquenta e quatro) folhas. Para efeitos práticos de explanação de como foi feita a análise, foram selecionadas duas sub-árvores. A primeira delas com nós “Faturamento”, “Segmento” e “Motivos específicos” para as empresas que não utilizam a Computação em Nuvem. Nesta configuração, apresentada na Figura 1, o classificador extraiu os melhores casos, identificados a partir da análise das folhas, aqui representadas por R13 a R19. Essa identificação da folha como “R” foi estabelecida para facilitar a interpretação dos resultados.

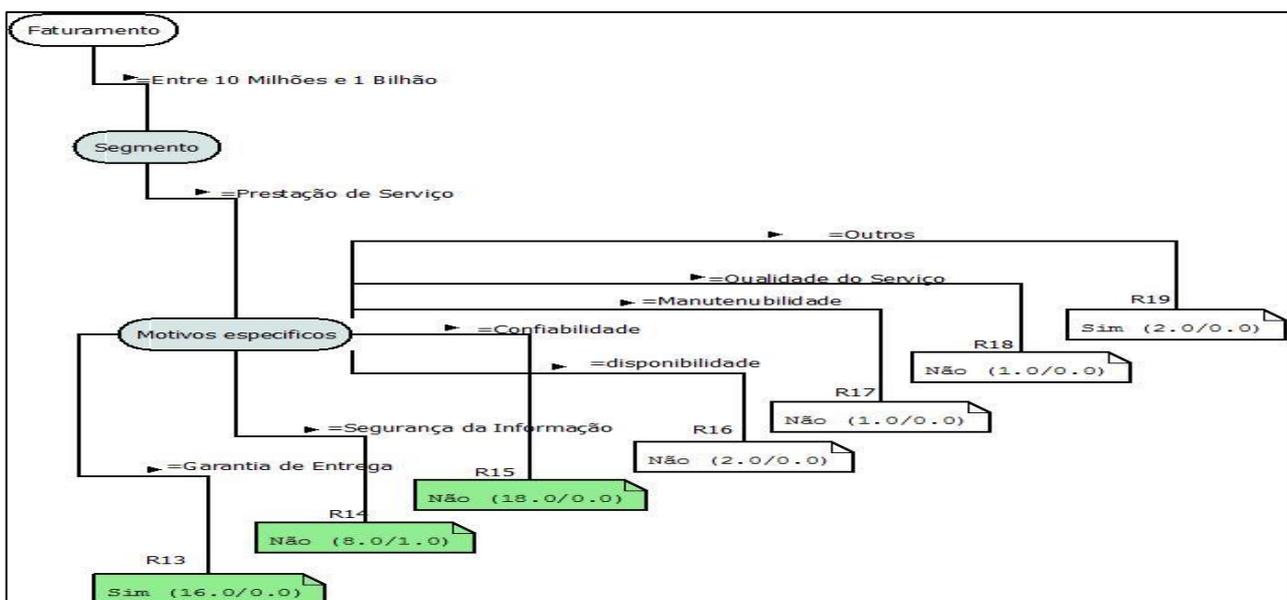


Figure 1: Árvore de decisão – Não utilizam Nuvem

- (R13) Empresas cujo faturamento anual se encontra “Entre 10 milhões e 1 Bilhão”, que são do segmento de “Comércio Varejista” e que possuem como motivo específico a “Garantia de Entrega”. O treinamento classificou

que contratariam os serviços de Computação em Nuvem, com base nas 16 (dezesseis) ocorrências do treinamento e 100% de acertos para essa classificação.

- (R14) Empresas cujo faturamento anual se encontra “Entre 10 milhões e 1 Bilhão”, que são do segmento de “Comércio Varejista” e que possuem motivo específico como “Segurança da Informação”. O treinamento classificou que a maioria não contrataria os serviços de Computação em Nuvem, com base nas 8 (oito) ocorrências do treinamento, sendo identificada apenas 1 (uma) instância classificada incorretamente.
- (R15) Empresas cujo faturamento anual se encontra “Entre 10 milhões e 1 Bilhão”, que são do segmento de “Comércio Varejista” e que possuem motivo específico como “Confiabilidade”. O treinamento classificou que não contratariam os serviços de Computação em Nuvem com base nas 18 (dezoito) ocorrências do treinamento, sendo identificado que 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.

A segunda sub-árvore apresenta os nós “Faturamento”, “Segmento” e “Características” para as empresas que não utilizam a Computação em Nuvem. Nesta configuração, apresentada na Figura 2, o classificador extrai os melhores casos, identificados a partir da análise das folhas, representadas por R10 a R13.

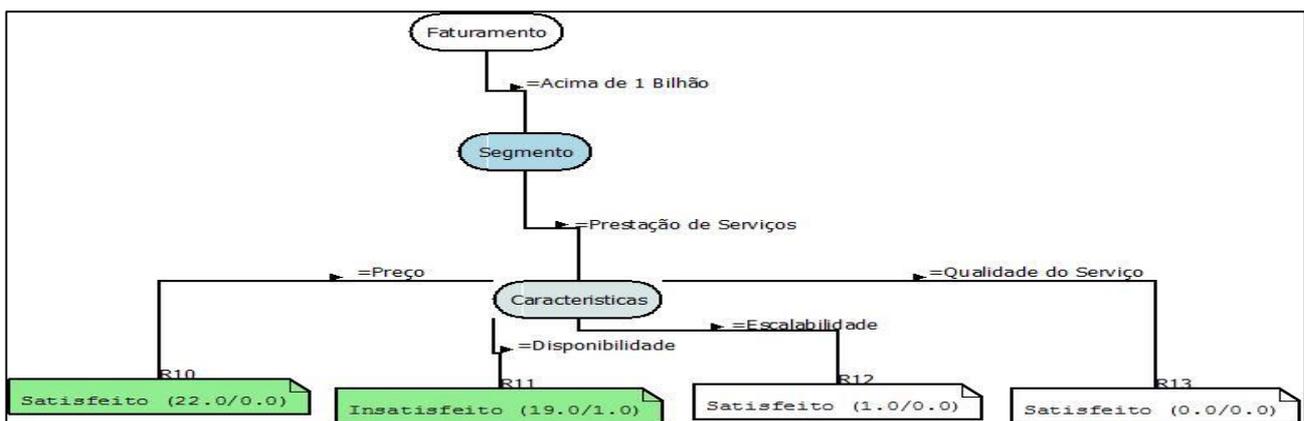


Figure 2: Árvore de decisão- Utilizam Nuvem.

- (R10) Empresas cujo faturamento anual se encontra “Acima de 1 bilhão”, que são do segmento de “Prestação de Serviços” e que possuem o preço como diferencial das “Características”. O treinamento classificou que estão satisfeitas com a contratação da Nuvem com base nas 22 (doze) ocorrências do treinamento, sendo identificado que 100% das instâncias foram classificadas corretamente.
- (R11) Empresas cujo faturamento anual é “Acima de 1 bilhão”, que são do segmento de “Prestação de Serviços” e que possuem a disponibilidade como diferencial das “Características”. O treinamento classificou que estão insatisfeitas com a contratação da nuvem com base nas 19 (nove) ocorrências do treinamento, sendo identificado que 0,52% das instâncias foram classificadas incorretamente.
- Para as classificações dos resultados R12 e R13, o classificador considerou os resultados insignificantes em virtude das ocorrências desse treinamento.

4.1 Descobertas e Discussão dos Resultados

A forma como foi elaborada a pesquisa, apresentando-se duas linhas de investigações, ocasionou a verificação de duas hipóteses: contratar ou não os serviços de Computação em Nuvem; e a satisfação ou insatisfação em relação aos serviços de Computação em Nuvem já contratados. Para esta verificação optou-se por selecionar na árvore de decisão apenas os casos cujo treinamento foi classificado com 100% de acertos. A Tabela 1 apresenta os casos relacionados à motivação para contratação ou não dos serviços de Computação em Nuvem, enquanto que a Tabela 2 se volta para os casos referentes à satisfação ou insatisfação dos serviços já contratados.

Treinamento 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

Tabela 1: Casos classificados com 100% de acertos

Foi descoberto que quando combinados os atributos de empresas que faturam entre 2 e 10 milhões atuando no Varejo, estas não estão propensas a contratarem os serviços de nuvem tendo como motivo a confiabilidade da informação (Sequência 1 da Tabela 1). Isto leva a prever que as empresas de médio porte ainda 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 de Computação Nuvem, existindo um forte indicativo que a garantia de entrega seja o motivo para essa tendência desses participantes. (Sequência 2 da Tabela 1).

Empresas que faturam um valor inferior a 500 mil por ano, que são de Outros segmentos e que possuem entre 21 (vinte e um) e 50 (cinquenta) colaboradores, descobriu-se que elas estão propensas a não contratar qualquer serviço de Computação em Nuvem (Sequência 3 da Tabela 1). Por se tratar de empresas de pequeno porte, elas não possuem recursos para tal investimento.

Outra descoberta é relativa a empresas que faturam entre 10 milhões e 1 bilhão e que são prestadoras de serviços (Sequência 4 da Tabela 1). O treinamento revelou que elas estão propensas a contratar os serviços de Computação em Nuvem devido a qualidade do serviço oferecido pelos provedores.

Treinamento classificados com 100% de acertos								
Seq	Treinamento	Combinação de Atributos			Folha	Instâncias	Classificador	
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

Tabela 2: Casos classificados com 100% de acertos (Satisfeito ou Insatisfeito)

Assim como nas descobertas anteriores, as presenças dos atributos de Faturamento e Segmento também são predominantes na verificação da segunda hipótese. A priori, descobriu-se que empresas com faturamento acima de 1 bilhão e que são prestadoras de serviços encontram-se satisfeitas com os serviços contratados de Computação em Nuvem tendo no preço seu principal motivo (Sequência 1 da Tabela 2).

Para outra descoberta, identificou-se que quando combinado faturamento entre 10 milhões e 1 bilhão e a utilização do modelo de nuvem pública, o motivo da grande satisfação dessas empresas é a contratação de Software como Serviço (Software as a Service – SaaS) (Sequência 2 da Tabela 2).

Por outro lado, para algumas empresas que possuem o mesmo faturamento e utilizam o modelo de nuvem pública, descobriu-se que a causa da insatisfação é a contratação de Banco de dados como Serviço (Database as a Service – DaaS) (Sequência 3 da Tabela 2).

Para a combinação de faturamento acima de 1 bilhão com o segmento de prestação de serviço, as empresas encontram-se insatisfeitas com a contratação dos serviços de Computação em Nuvem, sendo apontados a segurança e redução de custo como motivos dessa insatisfação (Sequência 4 da Tabela 2).

4.2 Ameaças aos Resultados

A pesquisa foi desenvolvida com variáveis diversificadas e categóricas através de um formulário *Web* e, portanto, é possível que alguns fatores possam ter influenciado nos resultados. Embora tenha sido realizado o Pré-Teste antes da liberação do *Survey* propriamente dito, algumas inconformidades foram detectadas na etapa de análise dos dados, o que pode ter prejudicado a interpretação dos resultados finais. Identificou-se que a não obrigatoriedade de algumas questões culminou em 140 (cento e quarenta) 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. É possível que o escopo da pesquisa que foi definido com 14 (quatorze) questionamentos possa não ter abrangido todas as áreas da Computação em Nuvem, como por exemplo, a existência de profissionais especializados para atuar na área de Computação em Nuvem, bem como a sua portabilidade como benefício que os provedores poderiam ofertar.

5 Considerações Finais

5.1 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 [37], 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 de Computação em 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? [37].

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 [37] 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: ao invés de se investigar o impacto dessa terceirização, abordaram-se aqui os motivos da não adesão/satisfação da Computação em Nuvem.

Outro trabalho relacionado foi um mapeamento sistemático foi realizado por [38]. Nessa pesquisa, o autor procurou realizar uma investigação na área de Computação em Nuvem sintetizando informações importantes. A proposta do autor foi que apresentar mecanismos que auxiliassem os interessados na adesão ao ambiente Nuvem, realizando para isso um mapeamento sistemático. 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? [37]

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.

Ainda para [38], 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.

5.2 Conclusões

A Computação em Nuvem está redefinindo a forma como as empresas passaram a se relacionar com a informação. É provável que a “dependência” dos profissionais pelo espaço/estrutura física esteja com seus dias contados, uma vez que a facilidade com que um ambiente de Computação em Nuvem pode disponibilizar esses recursos tornou mais prática, rápida e flexível às atividades das empresas.

Este trabalho apresentou um estudo realizado por meio de um *Survey* com o propósito de investigar os indícios pelos quais algumas empresas podem estar propensas a contratar ou não os serviços de Computação em Nuvem no Brasil. Adicionalmente, procurou-se identificar os motivos, para aquelas que já a usam, de sua satisfação ou insatisfação em relação aos serviços de Computação em Nuvem contratados no Brasil.

Em relação a primeira linha de investigação, pode-se evidenciar alguns indícios que tornam essas empresas propensas a contratar ou não os serviços de Computação em Nuvem. Primeiramente, destacam-se as empresas que estão propensas, em que os indícios estão relacionados ao seu alto faturamento e na aposta na garantia de entrega e qualidade dos serviços ofertados. Em contrapartida, as empresas que não estão propensas a contratar os serviços de Computação em Nuvem possuem como indícios o seu baixo faturamento e poucos colaboradores, associados à falta de confiança e segurança da informação nos serviços ofertados.

Em relação a outra linha de investigação da pesquisa, que é identificar os motivos de satisfação ou insatisfação dos contratantes dos serviços de Computação em Nuvem, pode-se destacar as descobertas dos motivos de satisfação, que referem-se ao preço do serviço associado aos modelos de Infraestrutura e Software como Serviço. Entretanto, descobriu-se que os motivos insatisfação estão associados a segurança, disponibilidade dos serviços e redução de custo.

A pesquisa revelou que a segurança influencia nos resultados tanto para tornar as empresas propensas à contratação, quanto nos motivos de insatisfação, pressupondo-se que essa variável tem papel negativo sob o ponto de vista das empresas que estão utilizando ou que pretendem utilizar.

Importante se ressaltar que 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. Em geral, tanto as descobertas dos aprendizados com 100% de acertos, como as que possuem a presença de erros, percebe-se a apresentação de resultados diversificados. Isto ocorre pois se entende que pesquisas com abordagem categórica tendem a apresentar tal problema. Logicamente, sabe-se da importância do aperfeiçoamento deste estudo, pois se deve proporcionar uma informação precisa e contundente para as comunidades profissional e acadêmica.

5.3 Trabalhos Futuros

Como propostas de trabalhos futuros e extensão 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.

Referências

- [1] Baun, C.; Kunze, M.; Nimis, J.; Tai, S. *Computação em Nuvem: web-based dynamic IT services*. Heidelberg, Berlin, Alemanha: Springer-Verlag, 2011.
- [2] Linthicum, D. S. *Computação em Nuvem and SOA Convergence in Your Enterprise: A Step-by-Step Guide*. 1. ed. São Paulo: Pearson Education, 2010. 265 p.
- [3] Linthicum, D. S. Sim, a nuvem está substituindo hardware e software corporativos Disponível em: <<http://cio.com.br/tecnologia/2013/12/03/sim-a-nuvem-esta-substituindo-hardware-e-software-corporativos/>> Acesso em 07 de janeiro de 2016.
- [4] Taurion, C. 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-os-tsunamis-transformadores-da-ti-para-os-ceos/>> Acesso em 14 de junho de 2015.
- [5] Bublitz, E. Catching the cloud: managing risk when utilizing Computação em Nuvem. *Property, Casualty Risk e Benefits Management*, 114(39), 2010, p. 12-16.
- [6] *Waikato Environment Knowledge Analysis – WEKA*. Disponível em: <https://www.weka-akademie.de/?chorid=00898157> Acesso em: 15 de novembro de 2015.
- [7] Mitchell, Tom M. *Machine Learning*. McGraw-Hill, 1997. Disponível em: <<http://www.cs.cmu.edu/afs/cs.cmu.edu/user/mitchell/ftp/tomhome.html>>. Acesso em: 11 de maio de 2016.
- [8] Tanenbaum, A. S., *Rede de Computadores*. 5. ed. Hardcover, 2010.
- [9] Taurion, C. Ameniza as tempestades de questionamentos sobre Computação em Nuvem. São Paulo, v. 1, p. 1, Novembro, 2011.
- [10] Mell, P; Grance T. The NIST Definition of Computação em Nuvem. Special Publication 800-145, 2009 Disponível em: <http://www.nist.gov/customcf/get_pdf.cfm?pub_id=909616> Acesso em 12 de dezembro de 2015.
- [11] Rittinghouse, J. W.; Ransome, F. J. *Computação em Nuvem: Implementation, Management and Security*. CRC PRESS, 2009.
- [12] Mohamed, A. A history of Cloud Computing. *ComputerWeekly.com*, 2009.

- [13] Conti, F. de. Grades computacionais para processamento de alto desempenho. Universidade Federal de Santa Maria, 2009, p. 3.
- [14] Alecrim, E. Cluster: conceito e características, InfoWester, 2013. Disponível em: <<http://www.infowester.com/cluster.php>> Acesso em 16-02-2016
- [15] Buyya, R. High Performance Cluster Computing: Architectures and Systems, Prentice Hall, volume 1. 1999.
- [16] Colvero, T. A; Dantas, M; Cunha, D. P., Ambientes de Clusters e Grids Computacionais: Características, Facilidades e Desafios. Santa Catarina, 2005.
- [17] Hurwitz, Judith et al. Cloud Computing For Dummies. Indianapolis: Wiley Publishing, 2010.
- [18] Foster, I.; Kesselman, C. The Grid 2: Blueprint for a New Computing Infrastructure. 2003.
- [19] Mahmood, Z; Hill, R. Computação em Nuvem for enterprise architectures. Londres, Inglaterra: Springer-Verlag, 2011.
- [20] Borges, H. P.; Sousa, J. M; Schulze, B; Mury, A. R. Computação em Nuvem. Brasil, 2011. Disponível em: <<http://livroaberto.ibict.br/handle/1/861>>. Acesso em: 16 de novembro de 2015.
- [21] Erl, T. Service Oriented Architecture: Concepts, Technology and Design. CrawsfordVille: Prentice Hall, 2005.
- [22] Russell, S.; Norvig, P. Artificial Intelligence: A Modern Approach. Prentice Hall, 2. ed, 2003.
- [23] Mohri, M. *Foundations of machine learning*. Cambridge, MA: MIT Press, 2012. p.16.
- [24] Camilo, C.O.; Carlos, J.S. Mineração de Dados: Conceitos, Tarefas, Métodos e Ferramentas. Universidade Federal de Goiás- UFG-RT-INF_001-09 - Relatório Técnico, 2009.
- [25] Breiman, L. Randomizing Outputs To Increase Prediction Accuracy. Technical Report 518, May 1, Statistics Department, UCB (in press Machine Learning), 1998.
- [26] Osório. F. Redes Neurais - Aprendizado Artificial , Fórum de Inteligência artificial, 1999.
- [27] Quinlan, J. R. C4.5- Programs for Machine Learning. Morgan Kauffman Publishers. San Mateo, CA. 1993. Disponível em: <<http://www.cs.su.oz.au/People/quinlan.html>>. Acesso em 10 de maio de 2016.
- [28] Nilsson, Nils J. *Artificial Intelligence: A New Synthesis*. Morgan Kaufmann Publishers. Disponível em: <<http://robotics.stanford.edu/people/nilsson/mlbook.html>> , 1998. Acesso em: 11 de maio de 2016.
- [29] Goldberg, D. E. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Publishing. 1989.
- [30] Revint, 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 [28] (REVINT, N° 1 2013).
- [31] Zuben, F J. V; Attux, R. F. Arvore de Decisão , Unicamp-SP, 2010.
- [32] Chapelle, O, Schölkopf, B., & Zien, A. (2006). Semi-supervised learning. MIT Press.
- [33] Stange, R. L; Neto, J. J. Aprendizagem Incremental usando Tabelas Decisão Adaptativas. 5° Workshop de Tecnologia Adaptativa – WTA, 2011.
- [34] Luger, G. F. Inteligência artificial: Estruturas e estratégias para a solução de problemas complexos. 4. ed. Bookman, Porto Alegre, 2007.
- [35] De Paula, M. B, Introdução automática de árvores de decisão. Pós-Graduação em Ciência da Computação, Universidade Federal de Santa Catarina-SC. Santa Catarina, 2002. 85 p.

[36] Zuben, F J. V; Attux, R. F. Arvore de Decisão , Unicamp-SP, 2010.

[37] Ramalho; C. N. L. R. Um Estudo sobre Adoção da Computação em Nuvem no Brasil. São Paulo: Universidade de São Paulo, 2012.

[38] Carvalho, F. J. S. Um mapeamento sistemático de estudos em Computação em Nuvem. Pernambuco: Universidade Federal de Pernambuco, 2012.