

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA  
E TECNOLOGIA FLUMINENSE

PROGRAMA DE PÓS-GRADUAÇÃO EM SISTEMAS APLICADOS À  
ENGENHARIA E GESTÃO

ANDERSON LUIZ IGNACIO DE LIMA

DESCOBERTA DE CONHECIMENTO ATRAVÉS DO ROTEIRO DE  
VISITA NA REGIÃO DO CRN4 COM MINERAÇÃO DE DADOS

Campos dos Goytacazes/RJ  
2020

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA  
E TECNOLOGIA FLUMINENSE

PROGRAMA DE PÓS-GRADUAÇÃO EM SISTEMAS APLICADOS À  
ENGENHARIA E GESTÃO

ANDERSON LUIZ IGNACIO DE LIMA

DESCOBERTA DE CONHECIMENTO ATRAVÉS DO ROTEIRO DE  
VISITA NA REGIÃO DO CRN4 COM MINERAÇÃO DE DADOS

Henrique Rego Monteiro da Hora  
(Orientador)

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação do Instituto Federal de Educação, Ciência e Tecnologia Fluminense, no Curso de Mestrado Profissional em Sistemas Aplicados à Engenharia e Gestão (MPSAEG), como parte dos requisitos necessários à obtenção do título de Mestre em Sistemas Aplicados à Engenharia e Gestão.

Campos dos Goytacazes/RJ  
2020

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA  
E TECNOLOGIA FLUMINENSE

PROGRAMA DE PÓS-GRADUAÇÃO EM SISTEMAS APLICADOS À  
ENGENHARIA E GESTÃO

ANDERSON LUIZ IGNACIO DE LIMA

DESCOBERTA DE CONHECIMENTO ATRAVÉS DO ROTEIRO DE  
VISITA NA REGIÃO DO CRN4 COM MINERAÇÃO DE DADOS

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação do Instituto Federal de Educação, Ciência e Tecnologia Fluminense, no Curso de Mestrado Profissional em Sistemas Aplicados à Engenharia e Gestão (MPSAEG), como parte dos requisitos necessários à obtenção do título de Mestre em Sistemas Aplicados à Engenharia e Gestão.

Projeto Apresentado em 28/02/2020.

Banca Examinadora:

---

Henrique Rego Monteiro da Hora  
Doutor em Engenharia de Produção - UFF  
(Orientador)

---

Sheila Andrade Abrahão  
Doutora em Ciência dos Alimentos - UFLA

---

Jonnathan dos Santos Carvalho  
Doutor em Computação - UFF

---

Helder Gomes Costa  
Doutor em Engenharia Mecânica – PUC-Rio

## LISTA DE FIGURAS

Figura 2.1 - Palavras-chave, tesouros e termos correspondentes.....	14
Figura 2.2 - Diagrama de <i>Venn</i> com a quantidade de trabalhos encontrados na pesquisa. ....	16
Figura 2.3 - Quantidade de ocorrências X Quantidade de veículos.....	17
Figura 2.4 - Quantidade de publicações por autores e coautores.....	17
Figura 2.5 - Artigos com maior número de citações.....	18
Figura 2.6 - Cronologia da produção. ....	19
Figura 2.7 - Distribuição das publicações entre os Países.....	20
Figura 2.8 - Árvore de palavras com as palavras encontradas nos resumos.....	21
Figura 3.1 - Fluxograma de procedimentos para busca e seleção dos artigos.....	30
Figura 3.2 - Diagrama de <i>Venn</i> com as técnicas utilizadas nas pesquisas selecionadas. ....	31
Figura 4.1 Etapas da estratégia de pesquisa. ....	42
Figura 4.2 - Árvore de Decisão - UAN - Desfecho Indicadores Qualitativos.....	52
Figura 4.3 - Árvore de Decisão - UAN - Desfecho Indicadores Quantitativos.....	53
Figura 4.4 - Árvore de Decisão – Alimentação Escolar - Desfecho Indicadores Qualitativos.....	55
Figura 4.5 - Árvore de Decisão – Alimentação Escolar - Desfecho Indicadores Quantitativos.....	56
Figura 4.6 - Árvore de Decisão – Hospitais e Instituições Similares - Desfecho Indicadores Qualitativos	58
Figura 4.7 - Árvore de Decisão - Hospitais e Instituições Similares - Desfecho Indicadores Quantitativos	59
Figura 4.8 - Árvore de Decisão - ILPI - Desfecho Indicadores Qualitativos.....	61
Figura 4.9 - Árvore de Decisão - ILPI - Desfecho Indicadores Quantitativos.....	62
Figura 4.10 - Quantidade de RVTs por ano.....	62

## LISTA DE QUADROS

Quadro 2.1 - Pesquisa em base de conhecimento. ....	14
Quadro 3.1 - Termos utilizados para busca. ....	29
Quadro 3.2 – Resumo das Informações dos estudos incluídos na revisão. ....	30
Quadro 4.1 - Atributos dos registros da base de dados coletada e suas respectivas descrições de conteúdo.....	44
Quadro 4.2 - Atributos e seus respectivos domínios da base de dados processada .....	45
Quadro 4.3 - Resumo dos domínios transformados com seus respectivos correlatos.....	46
Quadro 4.4 - Exemplo dos atributos conforme são utilizados nas fiscalizações .....	46
Quadro 4.5 - Parâmetros numéricos mínimos de referência para atuação do nutricionista - UAN. ....	47
Quadro 4.6 - Parâmetros numéricos mínimos de referência no âmbito do Programa de Alimentação Escolar (PAE). ....	48
Quadro 4.7 - Parâmetros numéricos mínimos de referência para atuação do nutricionista – Hospitais e Clinicas em Geral.....	48
Quadro 4.8 - Parâmetros numéricos mínimos de referência para atuação do nutricionista – ILPI.....	49
Quadro 4.9 - Matriz de confusão. ....	<b>Erro! Indicador não definido.</b>
Quadro 4.10 – Ganho de Informação - UAN. ....	51
Quadro 4.11 - Ganho de Informação - Alimentação Escolar. ....	54
Quadro 4.12 - Ganho de Informação - Hospitais e Instituições Similares. ....	57
Quadro 4.13 - Ganho de Informação - ILPI. ....	60
Quadro 4.14 - Valores de Acurácia em cada desfecho.....	63

## SUMÁRIO

1	APRESENTAÇÃO .....	8
1.1	Objetivos.....	9
2	ARTIGO 1 – FISCALIZAÇÃO DO EXERCÍCIO PROFISSIONAL COM APOIO DA INTELIGÊNCIA COMPUTACIONAL: BIBLIOMETRIA, TENDÊNCIAS E TRABALHOS MAIS RELEVANTES .....	10
2.1	Resumo .....	10
2.2	Abstract.....	10
2.3	Introdução .....	11
2.4	Metodologia .....	13
2.4.1	Definição do escopo da pesquisa .....	13
2.4.2	Pesquisa na amostra com uso de palavras-chave e seus tesouros .....	13
2.4.3	Identificação dos periódicos e conferências com maior número de artigos publicados .....	15
2.4.4	Identificação dos autores com maior número de publicações .....	15
2.4.5	Artigos com maior número de citações.....	15
2.4.6	Levantamento da cronologia da produção.....	15
2.4.7	Distribuição das publicações entre os Países.....	15
2.4.8	Identificação das palavras mais utilizadas.....	16
2.5	Resultados .....	16
2.5.1	Pesquisa na amostra com uso de palavras-chave e seus tesouros .....	16
2.5.2	Identificação dos periódicos com maior número de publicações .....	16
2.5.3	Identificação dos autores com maior número de publicações .....	17
2.5.4	Artigos com maior número de citações.....	18
2.5.5	Levantamento da cronologia da produção.....	19
2.5.6	Distribuição das publicações entre os Países.....	19
2.5.7	Árvore de palavras .....	20
2.6	Conclusão.....	21
2.7	Referências .....	23
3	ARTIGO 2 – MINERAÇÃO DE DADOS NA FISCALIZAÇÃO DO EXERCÍCIO PROFISSIONAL DA ÁREA DE SAÚDE: UMA REVISÃO SISTEMÁTICA. ....	26
3.1	Resumo .....	26
3.2	Abstract.....	26

3.3	Introdução .....	27
3.4	Metodologia .....	28
3.5	Resultados .....	29
3.6	Discussão .....	31
3.6.1	Mineração de dados para fiscalização do exercício profissional .....	32
3.6.2	Estatística como ferramenta de apoio ao exercício profissional .....	34
3.6.3	Estatística para fiscalização do exercício profissional .....	35
3.7	Considerações Finais .....	36
3.8	Referências .....	38
4	<b>ARTIGO 3 – MODELO CONCEITUAL DE ESTUDO DE FISCALIZAÇÃO PROFISSIONAL BASEADO EM MINERAÇÃO DE DADOS: UM ESTUDO DE CASO NO CONSELHO REGIONAL DE NUTRICINISTAS DA 4ª REGIÃO.</b> .....	39
4.1	Resumo .....	39
4.2	Abstract .....	39
4.3	Introdução .....	40
4.4	Mineração de Dados .....	41
4.5	Metodologia .....	42
4.5.1	Classificação da pesquisa .....	42
4.5.2	Definição e Coleta de Dados .....	43
4.5.3	Seleção, Pré-processamento e Transformação dos Dados .....	44
4.5.4	Procedimentos metodológicos e técnicas de análises (Mineração de Dados) .....	49
4.6	Análise e discussão dos resultados .....	50
4.6.1	Alimentação Coletiva - Unidade de Alimentação e Nutrição (UAN) .....	51
4.6.2	Alimentação Escolar (Gestor Público) .....	53
4.6.3	Nutrição Clínica - Hospital e Instituições Similares .....	56
4.6.4	Nutrição Clínica - Instituição de Longa Permanência para Idosos (ILPI) .....	59
4.6.5	Comparativo entre RVTs .....	62
4.7	Considerações Finais .....	64
4.8	Referências .....	66
5	<b>CONSIDERAÇÕES FINAIS</b> .....	69

## 1 APRESENTAÇÃO

Em rotinas de trabalho, há um grande volume de dados que são gerados e armazenados. Infelizmente, devido à incapacidade do ser humano em interpretar tamanha quantidade de dados, muita informação e conhecimento, possivelmente úteis, podem estar sendo desperdiçados, ficando ocultos dentro das Bases de Dados (Rezende, 2003). Esses dados gerados demandam técnicas e ferramentas que, com eficiência, potencializem a transformação dos dados em informação útil e oportuna (Carvalho et al., 2012).

Para potencializar a recuperação e o uso desses dados, uma das alternativas é o processo de Descoberta de Conhecimento em Bancos de Dados, do inglês *Knowledge Discovery in Database* (KDD) onde possui em uma das etapas a Mineração de Dados, também conhecido pelo termo em inglês *Data Mining*, o qual permite descobrir relações entre os dados armazenados mais facilmente do que com as técnicas tradicionalmente utilizadas (Carvalho et al., 2012).

Com a regulamentação de algumas profissões surgiu com elas a fiscalização profissional, que juntamente trouxe um grande volume de dados oriundos dessas fiscalizações. Esses dados são geralmente armazenados e requerem a utilização de técnicas que propiciem a interpretação e análise, com objetivo de agregar valor à atividade tanto do profissional fiscalizado quanto da rotina de fiscalização.

O conjunto de dados acumulados ao longo das fiscalizações realizadas compõe um rico material que é passível de tratamento além da estatística descritiva, e cuja exploração através das técnicas de Mineração de Dados pode expor padrões de comportamento que, caso conhecidos e expostos, tem grande potencial de utilização na gestão de fiscalizações e subsidiando futuras tomadas de decisões.

Esta pesquisa está estruturada em cinco capítulos. Este primeiro capítulo traz uma contextualização do tema. No segundo capítulo é apresentado um estudo bibliométrico de trabalhos que abordam os conceitos de Mineração de Dados, Fiscalização e Exercício Profissional. A partir dos trabalhos encontrados no primeiro artigo, e após identificado o surgimento de palavras relacionadas a área de saúde, foi realizado no terceiro capítulo uma segunda pesquisa incluindo os termos da área de saúde e em seguida, realizada uma análise sistemática, com o intuito de reunir estudos semelhantes, avaliando-os criticamente, o que deu origem ao segundo artigo. No quarto capítulo e conseqüente terceiro artigo, é realizado uma aplicação prática com o objetivo de gerar informações úteis a partir das bases de dados

disponíveis no sistema de informação do Conselho Regional de Nutricionistas da 4ª Região através de técnicas de Mineração de Dados, com intuito de expor padrões de comportamento das variáveis observadas. Finalizando, o quinto capítulo traz as considerações finais deste estudo, seguido pela lista de referências utilizadas na pesquisa.

### **1.1 Objetivos**

A presente pesquisa tem como objetivo geral a descoberta de conhecimento através do roteiro de visita na região do CRN4 com mineração de dados, e para isso analisar a produção científica publicada para obter um panorama sobre as pesquisas na área que une mineração de dados, fiscalização e exercício profissional. Em seguida, estabelecer um referencial teórico das etapas e métodos mais utilizados na área de mineração de dado na fiscalização do exercício profissional da área da saúde. E por fim, realizar uma aplicação prática buscando gerar informações úteis a partir das bases de dados disponíveis no sistema de informação do CRN4 através de técnicas de Mineração de Dados.

## 2 ARTIGO 1 – FISCALIZAÇÃO DO EXERCÍCIO PROFISSIONAL COM APOIO DA INTELIGÊNCIA COMPUTACIONAL: BIBLIOMETRIA, TENDÊNCIAS E TRABALHOS MAIS RELEVANTES

### 2.1 Resumo

Com a regulamentação de algumas profissões surgiu com elas a fiscalização profissional, que juntamente trouxe um grande volume de dados oriundos dessas fiscalizações. Esses dados são geralmente armazenados e requerem a utilização de técnicas que propiciem a interpretação e análise, com objetivo de agregar valor à atividade tanto do profissional fiscalizado quanto da rotina de fiscalização. O objetivo desta pesquisa é analisar os trabalhos publicados na área de inteligência computacional aplicados na fiscalização profissional, mais precisamente a utilização da mineração de dados como ferramenta para a fiscalização do exercício profissional. Como metodologia, é realizado um estudo bibliométrico na base de dados Scopus a partir da definição de conceitos chave com o intuito de analisar a produção acadêmica na área pesquisada. A partir das buscas combinando as três palavras chave definidas, são realizadas análises sobre principais autores, trabalhos mais citados, periódicos e conferências além de uma análise cronológica de produção. Os resultados demonstram uma baixa quantidade de publicações relacionadas a utilização de inteligência computacional aplicada a fiscalização profissional. Entretanto, a cronologia das produções evidencia que, apesar de ser uma área com um grande potencial a ser explorado, com publicações recentes onde a maioria dos trabalhos publicados são a partir de 2005, apresentam uma tendência de ascensão.

**Palavras-chave:** Mineração de dados; Fiscalização; Exercício profissional.

### 2.2 Abstract

*With the regulation of some professions emerged with them professional fiscalization, which together brought a large amount of data from these fiscalizations. These data are generally stored and require the use of techniques that provide interpretation and analysis, in order to add value to the activity of both the supervised professional and the inspection routine. The objective of this research is to analyze the published works in the area of computational intelligence applied in professional inspection, more precisely the use of data mining as a tool for the supervision of professional practice. As a methodology, a bibliometric study is carried*

*out in the Scopus database based on the definition of key concepts with the purpose of analyzing the academic production in the researched area. From the searches combining the three defined keywords, analyzes are carried out on the main authors, most cited works, periodicals and conferences, besides a chronological analysis of production. The results demonstrate a low number of publications related to the use of computational intelligence applied to professional supervision. However, the chronology of the productions shows that, despite being an area with great potential to be explored, with recent publications where most of the published works are from 2005, they show a tendency of ascension.*

**Keywords:** *Data mining; Fiscalization; Professional practice.*

### **2.3 Introdução**

As profissões são essencialmente a categoria de ocupação baseada no conhecimento, que geralmente segue um período de educação superior, formação e experiência profissional. Uma maneira diferente de categorizar essas ocupações é ver as profissões como os arranjos estruturais, ocupacionais e institucionais para lidar com o trabalho (Evetts, 2003).

Tem havido uma série de estudos que tentaram trabalhar dentro de abordagens interessantes e úteis para a compreensão do projeto profissional (Larson, 1977; Macdonald, 1995; Leicht, & Fennell, 2001) e a evolução contínua dos organismos profissionais. Muitos desses estudos se concentraram no estabelecimento dos corpos profissionais (Greenwood, Suddaby, & Hinings, 2002), sua relação com o Estado e seus esforços no processo de profissionalização, isto é, as ocupações se tornam aceitas como profissionais (Richardson, 1987).

Em muitas profissões, o conhecimento do funcionamento da regulamentação profissional é esperado ou mesmo exigido de todos os membros da profissão (Balthazard, 2017).

Existem muitas ocupações cujos membros afirmam ser profissionais. De fato, é difícil imaginar uma ocupação cujos membros não pretendem ser profissionais. Neste contexto, uma distinção importante é se a profissão é uma "profissão regulamentada" ou não. Ser reconhecido pelo estado como uma profissão regulamentada é visto por muitos como a linha de demarcação entre ser uma profissão "real" e ser apenas mais uma profissão auto-declarada. Para profissões estabelecidas, a regulamentação profissional é tida como certa; para profissões emergentes, no entanto, o que significa ser uma profissão regulamentada pode não ser tão

bem compreendido. A regulamentação profissional pode ser pensada como uma forma de proteção ao consumidor (Balthazard, 2017).

Para Posner (1975) e Noll (1989) a regulação profissional vem do interesse público em corrigir uma falha de mercado, onde baseia-se na ineficiência do equilíbrio de mercado. E a principal falha de mercado que se aplica aos mercados profissionais é a assimetria de informação (Stephen, & Love, 1999). Para a maioria dos clientes e consumidores, os serviços profissionais são produtos de credibilidade (Darby, & Karni, 1973).

Os benefícios da regulamentação incluem uma diminuição dos custos de pesquisa, melhorias na qualidade do serviço e fornecimento mais adequado de informações relativas à qualidade dos serviços profissionais. Além disso, e muito importante, uma redução no risco é esperada. De fato, devido à assimetria de informação, a regulação poderia ser o substituto mais adequado para o seguro (Garoupa, 2005).

Nesse contexto surgem os órgãos de fiscalização das atividades profissionais, que se revelam de extrema importância para o desenvolvimento sustentável na medida em que asseguram o adequado acompanhamento do exercício de atribuições previstas em lei, privativas de determinadas categorias profissionais. Esses órgãos possuem a finalidade de zelar pela integridade e pela disciplina das diversas profissões, disciplinando e fiscalizando, não só sob o aspecto normativo, mas também punitivo, o exercício das profissões regulamentadas, zelando pela ética no exercício destas. Cabe a estas entidades, além de defender a sociedade, impedir que ocorra o exercício ilegal da profissão, tanto por aquele que possua habilitação, mas não segue a conduta estabelecida, tanto para o leigo que exerce alguma profissão cujo exercício dependa de habilitação (Tribunal de Contas da União, 2014).

As fiscalizações são geralmente norteadas por roteiros ou formulários onde durante ou após as fiscalizações podem gerar um registro do observado, desta forma gerando uma grande quantidade de informações importantes para uma constante evolução das ações de fiscalização e do próprio exercício profissional. Dado o grande volume de dados, torna-se necessária a utilização de técnicas que contribuam para a análise deles. Para Machado, Nara, Schreiber & Schwingel (2015), bases de dados tornam-se pouco úteis sem a utilização de ferramentas de interpretação e análise.

Para potencializar o uso estratégico das informações, uma das alternativas é o processo de Descoberta de Conhecimento em Bancos de Dados, do inglês *Knowledge Discovery in Database* (KDD) onde possui em uma das etapas a Mineração de Dados, também conhecido

pelo termo em inglês *Data Mining*, o qual permite descobrir relações entre os dados armazenados mais facilmente do que com as técnicas tradicionalmente utilizadas (Carvalho et al., 2012).

O objetivo desta pesquisa é analisar os trabalhos publicados na área da mineração de dados de fiscalização com foco no exercício profissional, buscando verificar os autores mais citados, periódicos e os anos em que houveram mais publicações. Acredita-se que através desta pesquisa seja possível obter um panorama sobre as pesquisas na área que une mineração de dados, fiscalização e exercício profissional. Para Araújo Júnior, Perucchi, & Lopes (2013), a análise quantitativa da produção de uma área pode determinar tendências, e o conhecimento sobre essa produção aumenta a possibilidade de serem feitas inferências significativas para o desenvolvimento futuro.

## **2.4 Metodologia**

Com o intuito de identificar trabalhos que abordem a fiscalização do exercício profissional utilizando técnicas de mineração de dados é realizado um estudo bibliométrico. Para a elaboração desta pesquisa é utilizada como base a metodologia proposta por Costa (2010). São executadas 8 etapas descritas a seguir:

### **2.4.1 Definição do escopo da pesquisa**

Na primeira etapa, é definida como amostra os artigos indexados na base de conhecimento Scopus. Esta escolha se deve à representatividade e abrangência da base de dados que contém artigos de conferência, periódicos, anais, entre outros. Sendo que para compor o conjunto de trabalhos analisados, foram selecionados os artigos de periódicos e conferências. Em relação ao recorte temporal, a pesquisa contempla todos os trabalhos publicados até 2018.

### **2.4.2 Pesquisa na amostra com uso de palavras-chave e seus tesouros**

Para a segunda etapa, são definidas as palavras-chave, os tesouros e termos correspondentes para a realização da pesquisa na base Scopus.

Foram definidos, primeiramente, os conceitos que definem o método, o objeto de pesquisa e o objetivo: mineração de dados (método), exercício profissional (objeto de pesquisa) e fiscalização (objetivo). A partir daí foram definidas as palavras-chave: Mineração de dados (*Data Mining*), Fiscalização (*Fiscalization*), Exercício profissional (*Professional*

*Practice*). Além das palavras-chaves foram definidos seus tesouros e termos que também as representam na literatura, apresentados na Figura 2.1.

	Conceito A	Conceito B	Conceito C
	Mineração de dados	Fiscalização	Exercício profissional
Palavra-chave	Data Mining	Fiscalization	Professional Practice
Tesouros	Machine Learning	Inspection	Professional Standard
	Decision Tree	Audit	Professional Role
	SVM	Evaluation	Professional Council
	C45	Investigation	Professional Ethics
		Scrutinize	Professional Association
			Professional development
			Professionalism

**Figura 2.1 - Palavras-chave, tesouros e termos correspondentes.**

Como demonstrado na Figura 2.1, foi feita também uma organização classificando cada conceito como A, B e C. Essa classificação continuará a ser utilizada no decorrer deste trabalho.

Após a definição dos conceitos a serem utilizados foi formulada e realizada a pesquisa na base Scopus. A pesquisa inclui o corte de tipo de trabalho, considerando apenas os artigos de periódicos e de conferências. E o corte temporal, excluindo os trabalhos publicados a partir de 2019, por este ano não estar consolidado no momento da pesquisa, o que prejudicaria a série temporal. A pesquisa está apresentada no Quadro 2.1.

(TITLE-ABS-KEY ("data mining" OR "machine* Learning" OR "decision tree" OR "svm" OR "c45" )	#Tesouros de A
AND TITLE-ABS-KEY ("inspect*" OR "fiscaliz*" OR "audit*" OR "evaluat*" OR "investig*" OR "scrutin*" )	#Tesouros de B
AND TITLE-ABS-KEY ("Professional Practic*" OR "Professional Standard*" OR "Professional role" OR "Professional Council*" OR "Professional Ethics" OR "Professional Associat*" OR "Professional develop*" OR "Professionalism")	#Tesouros de C
AND ( LIMIT-TO ( DOCTYPE , "cp " ) OR LIMIT-TO ( DOCTYPE , " ar " ) )	#Corte de tipo
AND ( EXCLUDE ( PUBYEAR , 2019 ) )	#Corte temporal

**Quadro 2.1 - Pesquisa em base de conhecimento.**

Além da pesquisa apresentada no Quadro 2.1 também foram realizadas outras buscas na base Scopus, com 1 conceito ou combinando 2 conceitos. No total foram formuladas e executadas 7 pesquisas.

#### **2.4.3 Identificação dos periódicos e conferências com maior número de artigos publicados**

A partir dessa etapa, utilizou-se o pacote Bibliometrix, API Biblioshiny para fazer as avaliações dos artigos encontrados na pesquisa na base Scopus. O pacote Bibliometrix fornece um conjunto de ferramentas para pesquisa quantitativa em cientometria e bibliometria. (Aria, & Cuccurullo, 2017).

#### **2.4.4 Identificação dos autores com maior número de publicações**

Na terceira e quarta etapa, são identificados os periódicos e conferências além dos autores das publicações, adotando como recorte, os periódicos e os autores com duas ou mais publicações.

#### **2.4.5 Artigos com maior número de citações**

Na quinta etapa, é identificado os autores e os trabalhos com mais citações. São considerados os cinco trabalhos mais citados e são apresentados o título, os autores, o ano de publicação, o periódico ou conferência em que foi publicado e o número de citações.

#### **2.4.6 Levantamento da cronologia da produção**

Na sexta etapa, é realizado um estudo cronológico dos trabalhos produzidos e é apresentada a quantidade de publicações por ano em formato gráfico. Para esta etapa foram consideradas as pesquisas realizadas no Scopus que apresentavam os três conceitos definidos concomitantemente. O corte temporal foi realizado considerando as publicações até o ano de 2018.

#### **2.4.7 Distribuição das publicações entre os Países**

Na sétima etapa, é apresentado de forma gráfica a distribuição das publicações entre os Países, assim como a colaboração entre eles.

### 2.4.8 Identificação das palavras mais utilizadas

Na última etapa são identificadas as palavras mais utilizadas nos artigos, através das palavras definidas como palavras-chaves pelo autor, e as palavras mais utilizadas nos resumos. A representação é feita de forma visual através de uma árvore de palavras.

## 2.5 Resultados

### 2.5.1 Pesquisa na amostra com uso de palavras-chave e seus tesauros

O diagrama de *Venn* apresentado na Figura 2.2 apresenta a quantidade de trabalhos encontrados nas buscas realizadas na base Scopus, utilizando-se todas as combinações possíveis com os três conceitos definidos.



Figura 2.2 - Diagrama de *Venn* com a quantidade de trabalhos encontrados na pesquisa.

Na pesquisa que reuniu os conceitos mineração de dados e exercício profissional (AC) foram encontrados 118 trabalhos, considerado um número baixo se comparado à outras pesquisas que utilizaram dois conceitos, como, por exemplo, nas buscas dos conceitos AB e BC.

Uma quantidade ainda menor de trabalhos foi retornada na pesquisa que reuniu mineração de dados, fiscalização e exercício profissional, com o retorno de apenas 73 trabalhos, demonstrando que esta é uma área ainda pouco explorada.

### 2.5.2 Identificação dos periódicos com maior número de publicações

Nesta seção são apresentados os periódicos com mais publicações na pesquisa utilizando os três conceitos, onde foram retornados 62 periódicos e 11 anais de conferência.

Na Figura 2.3 é demonstrada a relação entre a quantidade de periódicos ou anais de eventos e a quantidade de publicações.



**Figura 2.3 - Quantidade de ocorrências X Quantidade de veículos.**

Através da Figura 2.3 é possível verificar que a maioria dos veículos retornados não possuem mais que dois trabalhos publicados. Para as pesquisas com os conceitos ABC foram retornados mais de 70 veículos com apenas uma publicação.

### 2.5.3 Identificação dos autores com maior número de publicações

Na terceira etapa da pesquisa são apresentados os autores com ao menos duas publicações.



**Figura 2.4 - Quantidade de publicações por autores e coautores.**

Na Figura 2.4 são apresentados 5 autores, pois foram os que apresentaram ao menos dois artigos publicados reunindo os conceitos mineração de dados, fiscalização e exercício profissional, todos os demais autores apresentaram apenas um artigo.

#### 2.5.4 Artigos com maior número de citações

Nesta etapa foram identificados os artigos mais citados na busca analisada. Para compor o gráfico da Figura 2.5 foram considerados os cinco artigos mais citados, classificados por ordem de quantidade de citações.

Analisando a Figura 2.5 é possível identificar que o artigo mais citado é um artigo relativamente recente, fruto de um estudo compartilhado. No quadro também pode ser observado que os artigos mais citados são majoritariamente relativos à área de saúde.

A opção por apresentar apenas 5 artigos foi devido à redução significativa em número de citações a partir do 6º colocado no ranking de citações, conforme pode ser visto no gráfico abaixo.

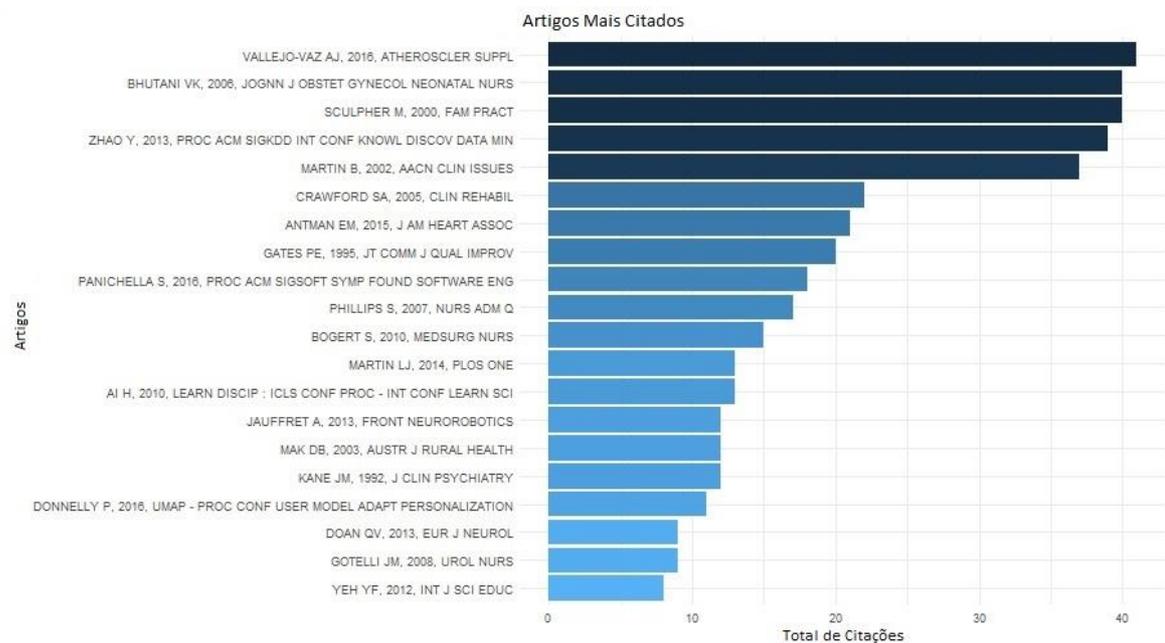
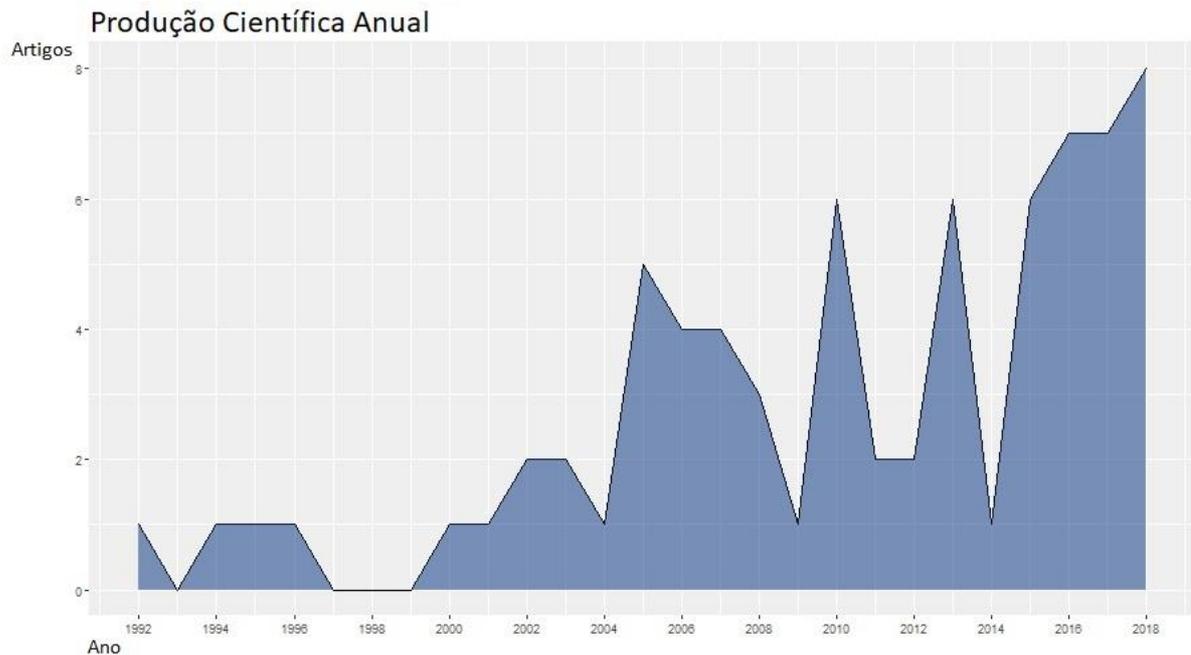


Figura 2.5 - Artigos com maior número de citações.

### 2.5.5 Levantamento da cronologia da produção

Para a análise da produção de trabalhos foram consideradas as pesquisas realizadas no Scopus que combinaram os três conceitos definidos. O corte temporal foi realizado considerando as publicações até o ano de 2018. A Figura 2.6 demonstra a cronologia da produção de 1992 a 2018. Nas buscas realizadas não foram retornados resultados publicações anteriores a 1992.



**Figura 2.6 - Cronologia da produção.**

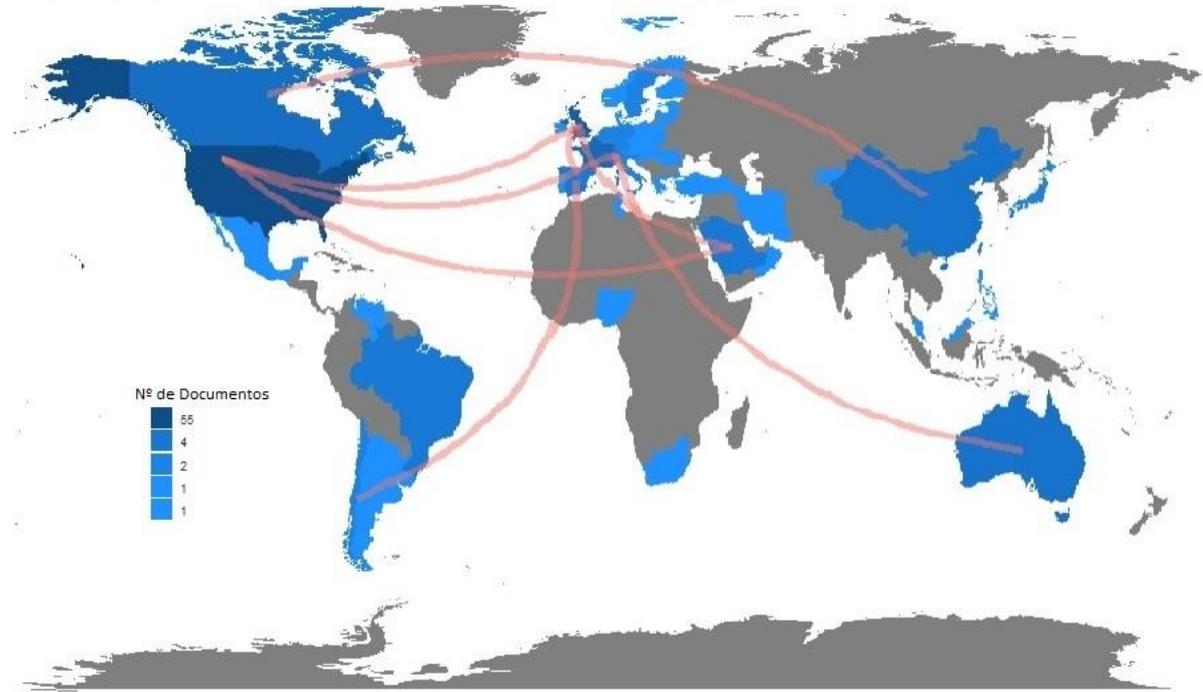
De acordo com o gráfico, os artigos encontrados a partir da pesquisa ABC que utilizam os conceitos mineração de dados, fiscalização e exercício profissional, começaram a ser publicados em 1992 e a ter um leve aumento em quantidade de publicações 2002.

De forma geral é possível observar que as publicações em sua maioria são posteriores a 2000, com aumento mais significativo a partir de 2005. O gráfico reforça a ideia de que na área de pesquisa que reúne mineração de dados, fiscalização e exercício profissional ainda há muito o que explorar.

### 2.5.6 Distribuição das publicações entre os Países

Nesta etapa foram identificados os países com o maior número de publicações, bem como a identificação da colaboração entre os países para a produção dessas publicações.

### Mapa de Produção e Colaboração Científica



**Figura 2.7 - Distribuição das publicações entre os Países.**

Observando o mapa acima é possível identificar que os países de língua inglesa são os que dominam tanto em quantidade de publicações quanto de colaboração.

Dentre os países com o maior número de publicações, os Estados Unidos estão em primeiro lugar com 55 publicações, seguido da França com 25 e em terceiro colocado fica o Reino Unido com 21. E entre os países mais colaborativos temos os Estados Unidos e o Reino Unido como os mais colaborativos.

Através da Figura 2.7, é possível observar que a Ásia tem pouco estudo sobre o tema, a Rússia não apresentou estudo, e o Brasil possui estudo, mas não colabora com outros países.

#### **2.5.7 Árvore de palavras**

A partir das palavras utilizadas nos resumos, foi montada a árvore de palavras abaixo. As árvores de palavras reúnem o conjunto de palavras mais utilizadas onde cada palavra é mostrada de modo que o tamanho da caixa está proporcionalmente de acordo com o número de vezes que a palavra aparece. Dessa forma é possível visualizar uma determinada palavra pela frequência da incidência.

data	clinical	care	study	program	based	management	approach	
professional	development	model	health	analysis	system	skills	decision	features
patients	learning	practice	patient	evaluation	methods	time	training	developed
				hospital	quality	implementation	social	

Figura 2.8 - Árvore de palavras com as palavras encontradas nos resumos.

A partir da árvore de palavras pode ser observado que além das palavras diretamente relacionadas à busca, como *data*, *professional* etc, surgem também palavras ligadas a área da saúde, como *clinical*, *care*, *patients*, *health*, isso nos mostra que as pesquisas estão mais direcionadas para esse campo de atuação que para outros. Essa frequência pode ser atribuída por se tratar de cuidado humano, onde sugere-se que é requerido uma maior fiscalização.

## 2.6 Conclusão

Esse estudo teve como objetivo central identificar trabalhos que abordem a utilização de técnicas de mineração de dados na fiscalização do exercício profissional e para tal foram feitas buscas na base Scopus.

Os dados da pesquisa relatados nesse artigo permitiram demonstrar que as pesquisas envolvendo mineração de dados, fiscalização e exercício profissional são recentes, com a maioria dos trabalhos publicados partir de 2005. Sendo uma área em ascendência, através do estudo bibliométrico, é notável que ainda é uma área pertinente e com muito a explorar.

As buscas com os três conceitos apresentaram resultados inferiores a 100, que pode ser considerado uma baixa quantidade, comparando-se buscas com apenas dois conceitos, com destaque para as buscas com os conceitos mineração de dados e fiscalização, onde apresentaram um resultado expressivo, com 93.550 resultados.

Em relação à análise sobre os autores com mais trabalhos publicados foi possível observar que nenhum deles possui uma quantidade considerável de publicações, sendo Blanchard N., D'Mello S.K., Kelly S., Nystrand M. e Olney A.M. os autores com maior

número de publicações e estes com apenas duas publicações cada, todos os demais apresentaram somente uma publicação.

Além de trazer um conhecimento sobre os trabalhos mais citados e autores com mais artigos publicados, este trabalho demonstra que a utilização de métodos de mineração de dados na fiscalização do exercício profissional é uma área ainda em exploração. Este fato pode ser notado principalmente através da análise cronológica onde é visível que a atenção a essa área começou a aumentar apenas a partir de 2005, porém ainda com oscilações em quantidade de publicações, se mantendo mais consistente somente a partir de 2015.

As principais limitações do estudo dizem respeito aos Tesouros utilizados, muito amplos para abarcar toda diversidade de autores e correntes podendo trazer resultados não relevantes ao campo de estudo. No entanto, os resultados aqui obtidos podem demonstrar as tendências e instrumentalizar pesquisas correlatas futuras.

Como trabalhos futuros propõe-se o aprofundamento em uma área de conhecimento, mais especificamente na área da saúde, visto que a partir da análise das nuvens de palavras e periódicos resultantes, os de maior frequência foram os dessa área.

## 2.7 Referências

- Araújo Júnior, R. H. de, Perucchi, V., & Lopes, P. R. D. (2013). Análise bibliométrica dos temas inteligência competitiva, gestão do conhecimento e conhecimento organizacional, no repositório institucional da universidade de Brasília. *Perspectivas em Ciência da Informação*, 18(4), 54–69. <https://doi.org/10.1590/S1413-99362013000400005>
- Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959–975. <https://doi.org/10.1016/j.joi.2017.08.007>
- Balthazard, C. (2017, setembro 1). *Professional Self-Regulation - HRP*. Human Resources Professionals Association.
- Barbosa, D. C. C., & Machado, M. A. (2007). Mineração de Dados usando o software WizRule em Base de Dados de Compras de TI. *Revista Eletrônica de Sistemas de Informação*, 6(1). <https://doi.org/10.21529/RESI.2007.0601001>
- Berry, M. J. A., & Linoff, G. (2004). *Data mining techniques: For marketing, sales, and customer relationship management* (2nd ed). Indianapolis, Ind: Wiley Pub.
- Berson, A., Smith, S., & Thearling, K. (2000). *Building data mining applications for CRM*. New York, NY: McGraw-Hill.
- Camilo, C. O., & Silva, J. C. da. (2009). *Mineração de Dados: Conceitos, Tarefas, Métodos e Ferramentas*. 29.
- Carvalho, D. R., Moser, A. D., Silva, V. A. da, & Dallagassa, M. R. (2012). Mineração de Dados aplicada à fisioterapia. *Fisioterapia em Movimento*, 25(3), 595–605. <https://doi.org/10.1590/S0103-51502012000300015>
- Costa, H. G. (2010). Modelo para webibliomining: proposta e caso de aplicação. *Revista da FAE*, (v.13, n.1), 115–126.
- Darby, M. R., & Karni, E. (1973). Free Competition and the Optimal Amount of Fraud. *Journal of Law and Economics*, p. 67–88.
- Evetts, J. (2003). The Sociological Analysis of Professionalism: Occupational Change in the Modern World. *International Sociology*, 18(2), 395–415. <https://doi.org/10.1177/0268580903018002005>
- Garoupa, N. M. (2005). Regulation of Professions in the US and Europe: A Comparative Analysis. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.640502>
- Gil, A. C. (2008). *Métodos e técnicas de pesquisa social*. São Paulo: Atlas.
- Greenwood, R., Suddaby, R., & Hinings, C. R. (2002). Theorizing change: The role of professional associations in the transformation of institutionalized fields. *Academy of Management Journal*, p. 58–80.
- Han, J., & Kamber, M. (2011). *Data mining: Concepts and techniques* (3rd ed). Burlington, MA: Elsevier.

- Hora, H. R. M. da, Shimoda, E., Erthal Júnior, M., & Azevedo Filho, E. T. (2018). Bibliometria da ética: Uma análise dos retrocessos de publicações em bases de conhecimento. *Pesquisa Brasileira em Ciência da Informação e Biblioteconomia*, 13(2). <https://doi.org/10.22478/ufpb.1981-0695.2018v13n2.42211>
- Larson, M. S. (1977). *The rise of professionalism: A sociological analysis*. Berkeley: University of California Press.
- Leicht, K. T., & Fennell, M. L. (2001). *Professional work: A sociological approach*. Malden, Mass., USA: Blackwell Publishers.
- Lobo, L. C. (2017). Inteligência Artificial e Medicina. *Revista Brasileira de Educação Médica*, 41(2), 185–193. <https://doi.org/10.1590/1981-52712015v41n2esp>
- Macdonald, K. M. (1995). *The sociology of the professions*. London ; Thousand Oaks, Calif: Sage.
- Machado, R. D., Nara, E. O. B., Schreiber, J. N. C., & Schwingel, G. A. (2015). ESTUDO BIBLIOMÉTRICO EM MINERAÇÃO DE DADOS E EVASÃO ESCOLAR. <https://doi.org/10.13140/rg.2.1.3164.3361>
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement*. 7.
- Naidus, E., & Celi, L. A. (2016). Big data in healthcare: Are we close to it? *Revista Brasileira de Terapia Intensiva*, 28(1). <https://doi.org/10.5935/0103-507X.20160008>
- Neves, B., Raimundo, A., & Obermeyer, Z. (2017). The Quiet Revolution of Big Data in Medicine. *Medicina Interna*, 24(4). <https://doi.org/10.24950/rspmi/Perspective/2017>
- Noll, R. G. (1989). Chapter 22 Economic perspectives on the politics of regulation. In *Handbook of Industrial Organization* (Vol. 2, p. 1253–1287). [https://doi.org/10.1016/S1573-448X\(89\)02010-8](https://doi.org/10.1016/S1573-448X(89)02010-8)
- Pina, M. de F., & Carvalho, M. S. (2017). GeoMed 2017: Visão mais profunda a partir de big data e pequenas áreas. *Cadernos de Saúde Pública*, 33(10). <https://doi.org/10.1590/0102-311x00172017>
- Posner, R. A. (1975). The Social Costs of Monopoly and Regulation. *University of Chicago Law School and National Bureau of Economic Research*, p. 807–827.
- Rezende, S. O. (Org.). (2003). *Sistemas inteligentes: Fundamentos e aplicações* (1. ed). Barueri, SP: Ed. Manole.
- Richardson, A. J. (1987). Professionalization and Intraprofessional Competition in the Canadian Accounting Profession. *Work and Occupations*, 14(4), 591–615. <https://doi.org/10.1177/0730888487014004006>
- Silva, E. L. da, & Menezes, E. M. (2005). *Metodologia da Pesquisa e Elaboração de Dissertação*. (4. ed. rev. atual.), 138.
- Stephen, F. H., & Love, J. H. (1999). Regulation of the legal profession. *University of Strathclyde*.

Tribunal de Contas da União. (2014). *Orientações para os Conselhos de Fiscalização das Atividades Profissionais*.

Weiss, S. M., & Indurkha, N. (1998). *Predictive data mining: A practical guide*. San Francisco: Morgan Kaufmann Publishers.

### 3 ARTIGO 2 – MINERAÇÃO DE DADOS NA FISCALIZAÇÃO DO EXERCÍCIO PROFISSIONAL DA ÁREA DE SAÚDE: UMA REVISÃO SISTEMÁTICA.

#### 3.1 Resumo

Os órgãos de fiscalização das atividades profissionais, em sua rotina de fiscalização, geram uma grande quantidade de informações importantes, que são importantes para uma constante evolução das ações de fiscalização e do próprio exercício profissional. Dado o grande volume de dados, torna-se necessário a utilização de técnicas que contribuam para sua análise. Esse trabalho visa estabelecer um referencial teórico das etapas e métodos mais utilizados na área de mineração de dado na fiscalização do exercício profissional da área da saúde, demonstrando as principais referências e autores do meio. A apresentação e estrutura da presente pesquisa foi baseada nos padrões da metodologia PRISMA (*Preferred Reporting Items for Systematic reviews and Meta-analyses*), através de pesquisas feitas nas bases Scopus, Web of Science, Science Direct e PubMed. A estratégia de busca nas bases de dados revelou 73 artigos no Scopus, 7 artigos no Web of Science, 1 artigo no Science Direct e 139 no PubMed. Dos 220 artigos, 37 foram excluídos na primeira triagem por estarem em duplicidade, em seguida é feito a exclusão após análise de elegibilidade, restando apenas 5 artigos. Destes, 2 utilizam mineração de dados ou texto para a fiscalização puramente para verificação da conduta do profissional, enquanto 2 utilizam a estatística com intuito de fiscalizar a fim de dar apoio a prática profissional, tendo o resultado da fiscalização como ferramenta para auxiliar o exercício profissional. E por último, apenas 1 artigo que utiliza a estatística como técnica unicamente para fiscalização da conduta profissional. Com este trabalho foi possível identificar a viabilidade de utilização de ferramentas matemáticas como forma de verificação ou apoio do exercício profissional, apesar de existirem poucos trabalhos na área.

**Palavras-chave:** Mineração de dados; Fiscalização; Exercício profissional; Saúde.

#### 3.2 Abstract

*The inspection bodies of professional activities, in their inspection routine, generate a large amount of important information, which is important for a constant evolution of inspection actions and professional practice itself. Given the large volume of data, it is necessary to use*

*techniques that contribute to its analysis. This work aims to establish a theoretical reference of the most used steps and methods in the area of data mining in the inspection of professional practice in the health area, demonstrating the main references and authors in the field. The presentation and structure of this research was based on the standards of the PRISMA methodology (Preferred Reporting Items for Systematic reviews and Meta-analyzes), through surveys carried out on Scopus, Web of Science, Science Direct and PubMed databases. The search strategy in the databases revealed 73 articles in Scopus, 7 articles in Web of Science, 1 article in Science Direct and 139 in PubMed. Of the 220 articles, 37 were excluded in the first screening because they were in duplicate, then the exclusion is made after analysis of eligibility, leaving only 5 articles. Of these, 2 use data or text mining for inspection purely to verify the professional's conduct, while 2 use statistics in order to monitor in order to support professional practice, with the result of inspection as a tool to assist professional practice . And finally, only 1 article that uses statistics as a technique only for the inspection of professional conduct. With this work it was possible to identify the feasibility of using mathematical tools as a way of verifying or supporting professional practice, although there are few studies in the area.*

**Keywords:** *Data mining; Fiscalization; Professional practice; Health.*

### **3.3 Introdução**

Os governos regulam uma grande quantidade de atividade comercial para garantir que o interesse público seja atendido. Uma dessas áreas são as transações entre profissionais e consumidores. Nesse contexto, a regulamentação profissional pode ser pensada como uma forma de proteção ao consumidor (Balthazard, 2017).

Os órgãos fiscalizadores possuem a finalidade de zelar pela integridade e pela disciplina das diversas profissões, disciplinando e fiscalizando, não só sob o aspecto normativo, mas também punitivo, o exercício das profissões regulamentadas, zelando pela ética no exercício destas. Cabe a estas entidades, além de defender a sociedade, impedir que ocorra o exercício ilegal da profissão, tanto por aquele que possua habilitação, mas não segue a conduta estabelecida, tanto para o leigo que exerce alguma profissão cujo exercício dependa de habilitação (Tribunal de Contas da União, 2014).

Com o início das fiscalizações, um grande volume de dados passou a ser gerado por

meio dessas fiscalizações. Esses dados são geralmente armazenados e requerem a utilização de técnicas que propiciem a interpretação e análise, com objetivo de agregar valor à atividade tanto do profissional fiscalizado quanto da rotina de fiscalização.

Uma alternativa para potencializar a recuperação e o uso desses dados, o qual permite descobrir relações entre os dados armazenados mais facilmente do que com as técnicas tradicionalmente utilizadas é a Mineração de Dados (Carvalho et al., 2012).

Uma área que emerge como destaque nessa questão de fiscalização profissional é a área da saúde. Apontada por Lima (2019) como sendo a área que teve mais frequência de aparições nas buscas sobre os temas de Mineração de Dados, Fiscalização e Exercício Profissional, e corroborando com o estudo de Hora et al. (2018), que em sua pesquisa sobre a ética na pesquisa, resultou como sendo a área que apresentou maior retrocessos de publicações. Esse destaque pode ser entendido por se tratar de uma área onde existe quantitativamente mais problemas éticos, e em função disso exige e demanda mais estudos de fiscalização do exercício profissional.

Nesse contexto, esse trabalho visa estabelecer um referencial teórico das etapas e métodos mais utilizados na área de mineração de dado na fiscalização do exercício profissional da área da saúde, demonstrando as principais referências e autores do meio.

### **3.4 Metodologia**

Utilizando-se como referência o estudo de Lima (2019) onde realizou-se pesquisa na base Scopus a partir das palavras-chave: *data mining*, *fiscalization* e *professional practice* e seus tesouros, incluiu-se os termos ligados a saúde a fim de identificar trabalhos da área da saúde que abordem a fiscalização do exercício profissional utilizando técnicas de mineração de dados, e ainda, ampliando a busca para outras bases. Os resultados encontrados formam a base de dados do presente trabalho.

A apresentação e estrutura da presente pesquisa foi baseada nos padrões da metodologia PRISMA (*Preferred Reporting Items for Systematic reviews and Meta-analyses*), proposta por Moher et al. (2009), descrita como uma técnica para apontar itens de relatório preferenciais para revisões sistemáticas e meta-análises.

As pesquisas foram feitas nas bases *Scopus*, *Web of Science*, *Science Direct* e *PubMed*. Sendo as três primeiras mais abrangentes e a última mais direcionada a área de biomedicina. A estratégia de busca utilizada na pesquisa é apresentada no Quadro 3.1.

**Quadro 3.1 - Termos utilizados para busca.**

	Conceito A	Conceito B	Conceito C	Conceito D
	Mineração de dados	Fiscalização	Exercício profissional	Saúde
Palavra-chave	Data Mining	Fiscalization	Professional Practice	Health
Tesauros	Machine Learning	Inspection	Professional Standard	Nourishment
	Decision Tree	Audit	Professional Role	Dietician
	SVM	Evaluation	Professional Council	Dietitian
	C45	Investigation	Professional Ethics	Nutrition
		Scrutinize	Professional Association	Care
			Professional development	Nurse
			Professionalism	Medical
				Clinic
				Patient

Após a busca, considerada no Quadro 3.1, foram excluídos os trabalhos duplicados, os restantes tiveram seus títulos e resumos avaliados quanto a exclusão, ou manutenção na análise. Para isto, foram utilizados como critérios de filtragem dos trabalhos: (i) ser da área de saúde, (ii) tratar da fiscalização ou verificação do exercício profissional, (iii) propunham ou utilizavam algum método para esse fim. Pesquisas foram excluídas quando: (i) não tratavam do exercício profissional; (ii) não apresentaram informação relevante.

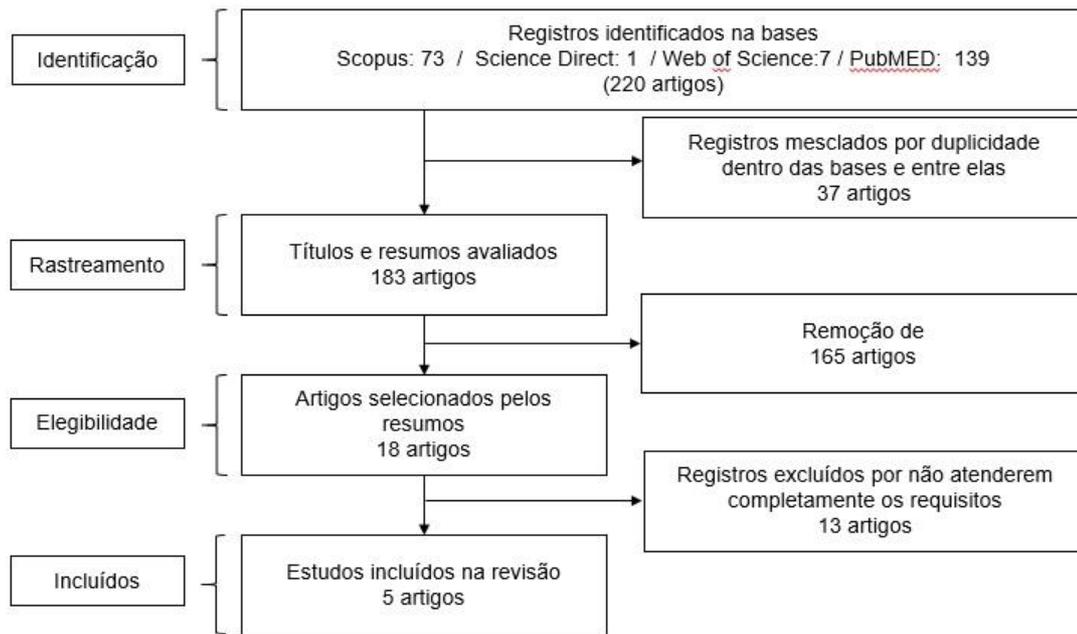
As referências obtidas na busca eletrônica foram armazenadas com o auxílio do programa Zotero versão 5.0.66, com o qual se procedeu à exclusão das duplicatas.

Os artigos extraídos da base de dados foram analisados, em uma primeira avaliação, tendo por base os títulos e o resumo dos artigos, e foram rejeitados aqueles que não preencheram os critérios de inclusão ou apresentaram algum dos critérios de exclusão. Quando um estudo não pôde ser incluído ou rejeitado com certeza, o texto completo foi analisado em uma segunda avaliação.

Por fim, é feito uma síntese dos estudos selecionados, que foram apresentados segundo suas características, e é feito uma descrição detalhada dos estudos propriamente ditos.

### 3.5 Resultados

A estratégia de busca nas bases de dados revelou 73 artigos no Scopus, 7 artigos no Web of Science, 1 artigo no Science Direct e 139 no PubMed. Dos 220 artigos, 37 foram excluídos na primeira triagem por estarem em duplicidade, em seguida é feita a exclusão após análise de elegibilidade. Ao final é apresentado os artigos que foram incluídos no estudo (Figura 3.1)



**Figura 3.1 - Fluxograma de procedimentos para busca e seleção dos artigos.**

Na sequência é identificada os principais métodos e técnicas que são utilizados na área de mineração de dados na fiscalização do exercício profissional da área da saúde, demonstrando as principais referências e autores do meio, além de verificar as conclusões mais assertivas sobre o tema em questão.

As pesquisas encontradas, assim como suas principais informações, são apresentadas pelo Quadro 3.2.

**Quadro 3.2 – Resumo das Informações dos estudos incluídos na revisão.**

Referência	Autores	Título	Periódico	Ano
A1	Guerrero, L.R., Ho, J., Christie, C., Harwood, E., Pfund, C., Seeman, T., McCreath, H., Wallace, S.P.	Using collaborative approaches with a multi-method, multi-site, multi-target intervention: Evaluating the National Research Mentoring Network	BMC Proceedings	2017
A2	Bernstein, D.N., Keswani, A., Ring, D.	Perioperative Risk Adjustment for Total Shoulder Arthroplasty: Are Simple Clinically Driven Models Sufficient?	Clinical Orthopaedics and Related Research	2017
A3	Gibbons, C., Richards, S., Valderas, J.M., Campbell, J.	Supervised machine learning algorithms can classify open-text feedback of doctor performance with human-level accuracy	Journal of Medical Internet Research	2017
A4	Hautemanière, A., Florentin, A., Hartemann, P., Hunter, P.R.	Identifying possible deaths associated with nosocomial infection in a hospital by data mining	American Journal of Infection Control	2011
A5	Layzell, M.	Improving the management of postoperative pain.	Nursing times	2005

Analisando o Quadro percebe-se que nenhum dos autores apareceram em mais de uma

das pesquisas e quanto aos periódicos em que foram publicados, nota-se que não há mais de uma publicação no mesmo periódico, o que demonstra uma pulverização nas publicações e indica que não há nem autor nem periódico especializados no tema.

Embora tenham em comum o objetivo de utilização de técnicas matemáticas/computacionais para fiscalização do exercício profissional, cada um dos trabalhos difere quanto a forma em que buscam esta finalidade e/ou a ação resultante da utilização destas técnicas. A Figura 3.2 demonstra como as pesquisas se segmentam em relação à técnica utilizada e a finalidade da fiscalização.



**Figura 3.2 - Diagrama de Venn com as técnicas utilizadas nas pesquisas selecionadas.**

A Figura 3.2 demonstra que existem duas segmentações quanto à técnica utilizada, se mineração de dados/texto ou estatística pura, e duas segmentações quanto aos objetivos, sendo a fiscalização unicamente com o objetivo de verificação do exercício profissional ou se o resultado da fiscalização serve de material de apoio a aprimoramento do próprio profissional.

A partir da figura 3.2, ainda é possível observar que dos 5 artigos, 2 utilizam mineração de dados ou texto para a fiscalização puramente para verificação da conduta do profissional, enquanto 2 utilizam a estatística com intuito de fiscalizar a fim de dar apoio a prática profissional, tendo o resultado da fiscalização como ferramenta para auxiliar o exercício profissional. E por último, apenas 1 artigo que utiliza a estatística como técnica unicamente para fiscalização da conduta profissional.

### **3.6 Discussão**

Nesta seção são apresentadas as intercessões ou agrupamentos que puderam ser

observados no diagrama de *Venn* apresentado na seção anterior. Foram avaliados os cinco trabalhos que apresentaram alguma ferramenta matemática para a fiscalização do exercício profissional dentro da área de saúde.

### **3.6.1 Mineração de dados para fiscalização do exercício profissional**

As pesquisas de Gibbons et al. (2017) e Hautemanière et al. (2011) objetivaram apresentar propostas para a fiscalização do exercício profissional, por intermédio da utilização da Mineração de Dados ou Mineração de Texto.

Gibbons et al. (2017) utilizam mineração de texto de feedback para avaliar variação de desempenho profissional dos médicos do Reino Unido, para isso utilizou comentários em texto aberto relativos ao desempenho de médicos coletados a partir de uma pesquisa com outros médicos, utilizando o *General Medical Council Colleague Questionnaire* (GMC-CQ). Os comentários foram codificados em 5 temas globais (inovação, habilidades interpessoais, popularidade, profissionalismo e respeito) usando uma estrutura qualitativa. Foram treinados 8 algoritmos de aprendizado de máquina para classificar comentários e avaliar seu desempenho usando várias amostras de treinamento. O desempenho do médico foi avaliado usando o GMC-CQ e as pontuações comparadas entre médicos com diferentes classificações usando testes t.

O objetivo do estudo foi treinar e avaliar um conjunto de algoritmos de aprendizado de máquina para classificar com precisão relatórios em texto aberto de médicos, que são conhecidos por serem positivamente tendenciosos, e avaliar o potencial de classificações baseadas em teoria em texto aberto para sinalizar diferenças no desempenho profissional dos médicos no Reino Unido.

Os resultados desta pesquisa apontaram que os algoritmos de aprendizado de máquina podem classificar o feedback de texto aberto do desempenho do médico, e que a capacidade das categorias em destacar diferenças no desempenho geral dos médicos são estatisticamente significativas.

Hautemanière et al. (2011) utilizam mineração de dados para identificar possíveis mortes por infecção hospitalar e para isso utilizou um banco de dados de mortalidade de um hospital universitário francês entre 1 de setembro de 2006 e 16 de setembro de 2007 com um total de 1.726 mortes registradas. Durante esse mesmo período, 6.290 potenciais infecções hospitalares foram identificadas por exame bacteriológico. Essas potenciais infecções hospitalares foram geradas usando um algoritmo de computador específico para o banco de dados de bacteriologia. Os formulários de solicitação de informações das potenciais infecções

hospitalares foram enviadas ao médico sênior da unidade onde as amostras foram obtidas para determinar se as potenciais infecções hospitalares eram uma infecções hospitalares, colonização ou outro tipo de infecção. Um total de 364 casos foram comuns aos dois bancos de dados; destes, uma amostra de 135 casos foi selecionada para análise posterior. Para estabelecer a força da evidência de infecção hospitalar como causa de morte, os 135 casos foram analisados usando o prontuário do paciente por um investigador da equipe do hospital.

O objetivo do presente estudo foi avaliar se vincular o banco de dados microbiológico ao banco de dados de mortalidade hospitalar pode ser uma abordagem de vigilância adequada para identificar pacientes que morreram com infecção hospitalar.

O estudo aponta uma estimativa de que a Infecção Hospitalar seja a principal causa de morte em 0,9% de todos os pacientes que morreram no hospital durante o período do estudo e uma causa contributiva em outros 8,0% desses pacientes, embora durante o período estudado não tenham sido relatados espontaneamente nenhuma morte associada à Infecção Hospitalar.

Esse resultado mostra que vincular bancos de dados da bacteriologia com aqueles que contêm registros de mortalidade hospitalar é uma ferramenta simples e reproduzível para identificar o número de mortes atribuíveis à Infecção Hospitalar. Isso pode fornecer uma abordagem poderosa para ajudar a reduzir o ônus da doença devido à Infecção Hospitalar, através da auditoria dessas mortes identificadas.

Ambos artigos apresentam grande relevância ao conteúdo trabalhado, pois fazem correlações importantes que podem auxiliar em pesquisas futuras utilizando as metodologias apresentadas, como a aplicação de algoritmos de aprendizado de máquina para classificar com precisão relatórios em texto aberto apresentado por Gibbons et al. (2017) e a utilização de bancos de dados como ferramenta simples e reproduzível para identificar o número de mortes atribuíveis à Infecções Hospitalares podendo ajudar a reduzir o ônus com doenças causadas por infecções hospitalares apresentado por Hautemanière et al. (2011).

Quanto a metodologia apresentada nos estudos, Gibbons et al. (2017) utilizaram como banco de dados o *General Medical Council Colleague Questionnaire* (GMC-CQ) com 1636 comentários em texto aberto (34.283 palavras) relativos ao desempenho de 548 médicos coletados a partir de uma pesquisa com colegas de médicos e foram treinados 8 algoritmos de aprendizado de máquina para classificar os comentários e avaliar seu desempenho usando várias amostras de treinamento. Em contrapartida, Hautemanière et al. (2011) utilizaram registros dos banco de dados de mortalidade de um hospital universitário francês e exames bacteriológico de duas bases de dados separadas, a de Potenciais Infecções Hospitalares e a de Mortalidade Hospitalar, ou seja, ambos utilizaram informações de acesso fácil e com

possibilidade de utilização em outros lugares que também possuem bancos de dados semelhantes.

Para a validação dos dados coletados, Gibbons et al. (2017) aplicaram, além da avaliação padrão do desempenho do algoritmo com um conjunto de dados de validação, a estabilidade do desempenho do algoritmo em diferentes dados também foi testada, usando uma validação cruzada de n-fold, diferente de Hautemanière et al. (2011) que empregaram para a classificação computacional do texto e a análise estatística o ambiente de programação estatística R (*R Foundation*) com o pacote “*RTextTools*” para treinar os algoritmos e o pacote “base” para realizar comparações entre grupos e para os dados foi utilizado o *software* de análise estatística SPSS versão 14.

Os resultados das análises apresentadas pelos dois artigos estudados foram positivos. Gibbons et al. (2017) apresentaram um alto desempenho do algoritmo individual, mesmo que a concordância entre os algoritmos e o codificador humano foi mais alta para uns códigos do que para outros.

No estudo de Hautemanière et al. (2011) não foram relatadas espontaneamente mortes associadas à Infecção Hospitalar. Dos 135 casos analisados, a Infecção Hospitalar foi considerada a principal causa de morte em 6 (4,4%) e fator contribuinte em 51 (37,8%). Assim, estimou-se que a Infecção Hospitalar seja a principal causa de morte em 0,9% de todos os pacientes que morreram no hospital durante o período do estudo e uma causa contributiva em outros 8,0% desses pacientes.

### **3.6.2 Estatística como ferramenta de apoio ao exercício profissional**

Os artigos de Guerrero et al. (2017) e Bernstein et al. (2017) utilizam estatística como ferramenta de apoio ao exercício profissional e por isso foram incluídas neste subgrupo.

Guerrero et al. (2017) utilizam a estatística para avaliar a Rede Nacional de Orientação para Pesquisa e com isso apoiar o sucesso da carreira de pesquisador da área biomédica. O artigo descreve o desenvolvimento inicial e a implementação do plano de avaliação colaborativa para a Rede Nacional de Pesquisa de Mentoreamento, cujos resultados são usados para entender como melhorar a orientação e desenvolvimento profissional para as minorias e, por sua vez, apoiar o sucesso da carreira de pesquisa biomédica sustentada.

A avaliação da Rede Nacional de Pesquisa de Mentoreamento emprega um processo de múltiplos e mistos processo e desenho de resultado que inclui pesquisas de avaliação padrão, observações e métodos quase-experimentais.

Os resultados da avaliação são utilizados para informar a política nacional sobre a

melhor forma de conceber e implementar iniciativas abrangentes de orientação e desenvolvimento de carreira para diversas populações que buscam carreiras em pesquisa biomédica.

Bernstein et al. (2017) utilizam estatística para comparar os resultados de modelo estatístico versus avaliação clínica feita por especialistas para casos de readmissão não planejada dentro de 30 dias após alta de pacientes submetidos a artroplastia total do ombro (TSA) e desde modo estimar o risco, além da intuição clínica.

Os autores construíram um modelo de regressão logística para evento adverso e readmissão não planejada em 30 dias usando cinco variáveis escolhidas a priori com base na experiência clínica.

Como resultado obteve-se que o modelo estatisticamente conduzido teve melhor desempenho e explicou melhor a variação na readmissão não planejada dentro de 30 dias após a alta após uma TSA em comparação com o modelo dirigido clinicamente.

Deste modo, comparando o modelo clinicamente conduzido com modelos de risco estatisticamente derivados de artroplastia total do ombro (TSA) oferece indícios sobre possíveis lacunas entre a prática comum e medicina baseada em evidências.

Relacionando as duas pesquisas mencionadas nesta subseção, ambas utilizam estatística, porém de formas diferentes, Guerrero et al. (2017) a utilizam para avaliar eficácia e alcance do Programa da Rede Nacional de Pesquisa de Mentoreamento, em contrapartida, Bernstein et al. (2017) utilizam estatística para comparar os resultados de modelo estatístico versus avaliação clínica feita por especialistas. Deste modo o primeiro avalia eficácia e o segundo estima riscos.

Do ponto de vista metodológico, Guerrero et al. (2017) coletam as informações através do portal da própria Rede de Pesquisa e do sistema interno de registro, por outro lado, Bernstein et al. (2017) utilizam dados do Programa Nacional de Melhoria da Qualidade Cirúrgica do Colégio Americano de Cirurgiões do período de 2011 a 2014, ambas são bases específicas, porém disponíveis aos pesquisadores.

Observou-se também que os estudos traçaram caminhos diferentes para as análises, Guerrero et al. (2017) tiveram como alvo os profissionais, enquanto Bernstein et al. (2017) utilizaram os pacientes.

### **3.6.3 Estatística para fiscalização do exercício profissional**

Layzell (2005) contextualiza sobre prescrição inadequada de medicamentos em pacientes pós-operatório, que objetivo identificar, através de auditorias, os atrasos sofridos

pelos pacientes para o alívio da dor no pós-operatório e os motivos.

As auditorias mostraram que altas proporções de pacientes estão acordando da anestesia geral na dor e um número substancial está na dor moderada a intensa. Isto significa que a prescrição inadequada pelos anestesistas causa sofrimento desnecessário e efeitos potencialmente prejudiciais aos pacientes.

Apesar da população atendida em geral apresentar alto grau de confiança na capacidade de enfermeiros e médicos para tratar sua dor e seus conhecimentos e habilidades no gerenciamento da dor serem importantes no cuidado com o paciente com dor, atitudes e crenças inadequadas entre médicos e enfermeiros demonstraram afetar o controle da dor. Além disso, existem muitas barreiras dos pacientes ao bom manejo da dor que devem ser reconhecidas e abordadas pela equipe, além do "déficit de conhecimento" dos profissionais que pode ser a causa mais prevalente do manejo inadequado da dor, com prescrições precárias por médicos e administração inadequada por enfermeiros.

Diante desse contexto foi realizada uma auditoria para identificar os atrasos sofridos pelos pacientes e seus motivos.

Os dados foram coletados na enfermaria de recuperação de pós-operatório por um período de cinco semanas. Todos os pacientes adultos no pós-operatório relataram um escore de dor verbal de três ou mais ao acordar (escala de 0 a 10, onde 0=sem dor e 10=dor insuportável) e após análise estatística os resultados encontrados na auditoria mostrou que altas proporções de pacientes estão acordando da anestesia geral na dor e um número substancial está na dor moderada a intensa.

A partir desse resultado estatístico obtido após auditoria foi possível desenvolver um novo sistema usando prescrições padrão para capacitar os enfermeiros a gerenciar a dor dos pacientes de maneira rápida e eficaz. Os dados da auditoria mostram claramente que, se a equipe de recuperação tiver prescrições adequadas para agir rapidamente sobre a dor do paciente, o tempo necessário para atingir uma pontuação razoável de dor será reduzido pela metade.

### **3.7 Considerações Finais**

O presente trabalho teve como propósito conduzir uma revisão sistemática da literatura sobre os métodos mais utilizados na área de mineração de dados na fiscalização do exercício profissional, exclusivamente da área da saúde, pois esta foi a área que, segundo Lima (2019) apareceu com mais frequência nas buscas sobre os temas abordados, além ainda de ter sido apontada por Hora et al. (2018) como sendo a área onde tem mais problemas éticos em

publicações, então, pode-se ter uma suposição ou levanta-se a hipótese de que se deva ter também problemas éticos no exercício profissional. E por ter mais problemas éticos nas publicações e por hipótese mais problemas éticos no exercício profissional é onde demanda mais fiscalização ou verificação, e onde as ferramentas de apoio computacional poderiam ser mais bem utilizadas, mas isso ainda não acontece.

Através desta pesquisa, foi possível identificar que existem poucos trabalhos na área, deste modo há ainda um campo importante e amplo a ser estudado. As ferramentas de computação podem estar mais fundidas em outras áreas de conhecimento, mas não ainda na saúde.

Gibbons et al. (2017) e Hautemanière et al. (2011) concluíram a aplicabilidade dos métodos apresentados, onde os algoritmos de aprendizado de máquina podem classificar o feedback de texto aberto do desempenho do médico em vários temas derivados de avaliadores humanos com alto desempenho, mesmo quando há tempo e recursos limitados para codificar conjuntos de dados de treinamento. Essas descobertas podem informar futuros modelos preditivos de desempenho e apoiar a avaliação em tempo real para melhorar a qualidade e a segurança. Demonstrou ainda que, ao vincular dois bancos de dados existentes, é possível identificar óbitos relacionados à Infecções hospitalares em pacientes hospitalizados, sendo uma ferramenta muito útil para identificar essas mortes para fins de auditoria e de grande valia na redução do ônus da doença devido a esse problema de saúde pública.

Como resultados, Guerrero et al. (2017) proporcionou um estreito relacionamento de trabalho colaborativo das equipes de avaliação, de modo a produzir continuamente planos fortes e claros para processos de avaliação coordenados que alavanquem de maneira eficiente e efetiva os recursos. Bernstein et al. (2017) mostrou que o modelo estatisticamente conduzido apresentou melhor desempenho do que o modelo clinicamente orientado, deste modo serve de auxílio na melhoria da qualidade cirúrgica e redução de custos, ressaltando a importância de abordagens adequadas de estratificação de risco na seleção e no manejo de pacientes.

Layzell (2005) concluiu que o tempo economizado pela equipe de enfermagem de recuperação e enfermagem com essa abordagem simples de implementar uma etiqueta de prescrição padrão deve melhorar a experiência pós-operatória para a maioria das admissões cirúrgicas de rotina e a equipe poderá se dedicar a outros cuidados aos pacientes.

### 3.8 Referências

- Balthazard, C. (2017, setembro 1). *Professional Self-Regulation - HRP*. Human Resources Professionals Association.
- Bernstein, D.N., Keswani, A., Ring, D. (2017). Perioperative Risk Adjustment for Total Shoulder Arthroplasty: Are Simple Clinically Driven Models Sufficient?. *Clinical Orthopaedics and Related Research*. <https://doi.org/10.1007/s11999-016-5147-y>
- Carvalho, D. R., Moser, A. D., Silva, V. A. da, & Dallagassa, M. R. (2012). Mineração de Dados aplicada à fisioterapia. *Fisioterapia em Movimento*, 25(3), 595–605. <https://doi.org/10.1590/S0103-51502012000300015>
- Gibbons, C., Richards, S., Valderas, J.M., Campbell, J. (2017). Supervised machine learning algorithms can classify open-text feedback of doctor performance with human-level accuracy. *Journal of Medical Internet Research*. <https://doi.org/10.2196/jmir.6533>
- Guerrero, L.R., Ho, J., Christie, C., Harwood, E., Pfund, C., Seeman, T., McCreath, H., Wallace, S.P. (2017). Using collaborative approaches with a multi-method, multi-site, multi-target intervention: Evaluating the National Research Mentoring Network. *BMC Proceedings*. <https://doi.org/10.1186/s12919-017-0085-6>
- Hautemanière, A., Florentin, A., Hartemann, P., Hunter, P.R. (2011). Identifying possible deaths associated with nosocomial infection in a hospital by data mining. *American Journal of Infection Control*. <https://doi.org/10.1016/j.ajic.2010.04.216>
- Hora, H. R. M. da, Shimoda, E., Erthal Júnior, M., & Azevedo Filho, E. T. (2018). Bibliometria da ética: Uma análise dos retrocessos de publicações em bases de conhecimento. *Pesquisa Brasileira em Ciência da Informação e Biblioteconomia*, 13(2). <https://doi.org/10.22478/ufpb.1981-0695.2018v13n2.42211>
- Layzell, M. (2005). Improving the management of postoperative pain. *Nursing times*. 2005 Jun 28-Jul 4;101(26):34-6.
- Lima, A. L. I. de (2019). Fiscalização do Exercício Profissional com Apoio da Inteligência Computacional: Bibliometria, Tendências e Trabalhos mais Relevantes.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement*. 7.
- Tribunal de Contas da União. (2014). *Orientações para os Conselhos de Fiscalização das Atividades Profissionais*.

#### 4 ARTIGO 3 – MODELO CONCEITUAL DE ESTUDO DE FISCALIZAÇÃO PROFISSIONAL BASEADO EM MINERAÇÃO DE DADOS: UM ESTUDO DE CASO NO CONSELHO REGIONAL DE NUTRICINISTAS DA 4ª REGIÃO.

##### 4.1 Resumo

Dada a problemática da quantidade de dados acumulados ao longo das fiscalizações realizadas pelos Conselhos Profissionais, através da utilização da mineração de dados é possível expor padrões de comportamento das empresas e profissionais fiscalizados que, caso conhecidos e expostos, teriam grande potencial de utilização na gestão de fiscalizações e subsidiando futuras tomadas de decisões. Nesse cenário, objetivo do presente trabalho é gerar informações úteis a partir das bases de dados através de técnicas de Mineração de Dados, com intuito de expor padrões de comportamento das variáveis observadas. Para o cumprimento do objetivo é realizado um estudo de caso no Conselho Regional de Nutricionistas da 4ª Região. São executadas as etapas de Definição e Coleta de Dados; Seleção, Pré-processamento, higienização e Transformação dos Dados; Mineração de Dados e Análise e Discussão dos Resultados. Na última etapa é utilizado o método de classificação J48. A partir das árvores geradas pelo método foi possível concluir que é possível encontrar uma tendência de resultado e possíveis padrões nos desfechos propostos para o processo de mineração de dados utilizando-se a base de dados oriunda de fiscalizações anteriores.

**Palavras-chave:** Mineração de dados; Fiscalização; Exercício profissional; Nutrição.

##### 4.2 Abstract

*Given the problem of the amount of data accumulated over the inspections carried out by the Professional Councils, through the use of data mining they can expose behavior patterns of the companies and supervised professionals that, if known and exposed, have a great potential for use in the management of inspections and subsidizing future decision-making. In this scenario, the objective of the present paper is to generate useful information from the databases through Data Mining techniques, in order to expose behavior patterns of the observed variables. To achieve the objective, a case study is carried out at the Regional Council of Nutritionists of the 4th Region. The steps performed are Data Definition and*

*Collection; Data Selection, Pre-processing, Sanitization and Transformation; Data Mining and Analysis and Discussion of Results. In the last step, the J48 classification method is used. From the trees generated by the method, it was possible to conclude that it is possible to find a trend of results and possible patterns in the proposed outcomes for the data mining process using the database from previous inspections.*

**Key words:** *Data mining; Inspection; Professional practice; Nutrition.*

### **4.3 Introdução**

Uma sociedade provida por tecnologia da informação é capaz de produzir uma elevada quantidade de dados, cujo relevância cresce no mesmo ritmo exponencial (Neves, Raimundo, & Obermeyer, 2017), e não é diferente nas áreas correlatas à saúde coletiva (Pina & Carvalho, 2017). A disponibilidade destes grandes volumes de dados, aliado à aplicações de técnicas de inteligência artificial são capazes de propor soluções aos problemas de saúde, até então não possíveis (Lobo, 2017).

Nesse contexto alguns estudos apresentam a utilização de ferramentas matemáticas para o tratamento de dados dentro da área da saúde e relatam a obtenção de resultados satisfatórios. Bernstein et al. (2017) utilizam estatística para comparar os resultados de modelo estatístico versus avaliação clínica feita por especialistas para casos de readmissão não planejada dentro de 30 dias após alta de pacientes e desde modo estimar o risco, além da intuição clínica. Para os autores, o modelo estatisticamente conduzido apresentou melhor desempenho do que o modelo clinicamente orientado, servindo de auxílio na melhoria da qualidade cirúrgica e redução de custos.

Gibbons et al. (2017) e Hautemanière et al. (2011) utilizam técnicas de mineração para fiscalização do exercício profissional. Gibbons et al. (2017) utilizam mineração de texto de feedback para avaliar variação de desempenho profissional dos médicos do Reino Unido, e Hautemanière et al. (2011) utilizam mineração de dados para identificar possíveis mortes por infecção hospitalar. Ambos concluíram a aplicabilidade dos métodos apresentados, indicando que suas descobertas podem informar futuros modelos preditivos de desempenho e apoiar a avaliação para melhorar a qualidade do exercício profissional.

Algumas das finalidades do Conselho Regional de Nutricionistas é Orientar, supervisionar, fiscalizar e disciplinar o exercício da profissão, buscando a qualidade dos serviços relacionadas à alimentação e nutrição (Lei Nº 6.583, 1978). Nas ações de fiscalização

uma das ferramentas utilizadas pelas fiscais são os roteiros de visita técnica que foram elaborados baseado na resolução que dispõe sobre a definição das áreas de atuação do nutricionista e suas atribuições (Resolução CFN Nº 380, 2005), esses roteiros estão disponíveis em <http://www.crn4.org.br/fiscalizacao-roteiro-de-visita-tecnica.php>, onde os dados gerados a partir das visitas são inseridos em uma plataforma própria.

O conjunto de dados acumulados ao longo das fiscalizações realizadas compõe um rico material que é passível de tratamento além da estatística descritiva, e cuja exploração através das técnicas de Mineração de Dados podem expor padrões de comportamento das empresas e profissionais fiscalizados que, caso conhecidos e expostos, tem grande potencial de utilização na gestão de fiscalizações e subsidiando futuras tomadas de decisões.

Nesse cenário, este artigo tem por objetivo gerar informações úteis a partir das bases de dados disponíveis no sistema de informação do Conselho Regional de Nutricionistas da 4ª Região através de técnicas de Mineração de Dados, com intuito de expor padrões de comportamento das variáveis observadas.

#### **4.4 Mineração de Dados**

Muito conhecimento encontra-se escondido na vasta quantidade de dados disponíveis nos bancos de dados das empresas e é com a Mineração de Dados que se pode transformar esses dados brutos em informação valiosa a fim de auxiliar a tomada de decisão (Barbosa, & Machado, 2007).

Segundo Berson, Smith, & Thearling (2000), a Mineração de Dados explora as bases de dados através de dezenas de centenas de pontos de vista diferentes. A diferença entre a Mineração de Dados e as Técnicas Estatísticas está na utilização dos próprios dados para a descoberta dos padrões e não na verificação de padrões hipotéticos.

As bases de dados armazenam conhecimento que podem auxiliar na melhoria dos negócios, e as técnicas tradicionais permitem somente verificar hipóteses que são, aproximadamente, apenas 5% de todas as relações encontradas por esses métodos. A Mineração de Dados pode descobrir as outras relações desconhecidas: os 95% restantes (Barbosa, & Machado, 2007).

Sendo assim, Mineração de Dados é um conjunto de técnicas que envolvem métodos matemáticos, algoritmos e heurísticas para descobrir padrões e regularidades em grandes conjuntos de dados.

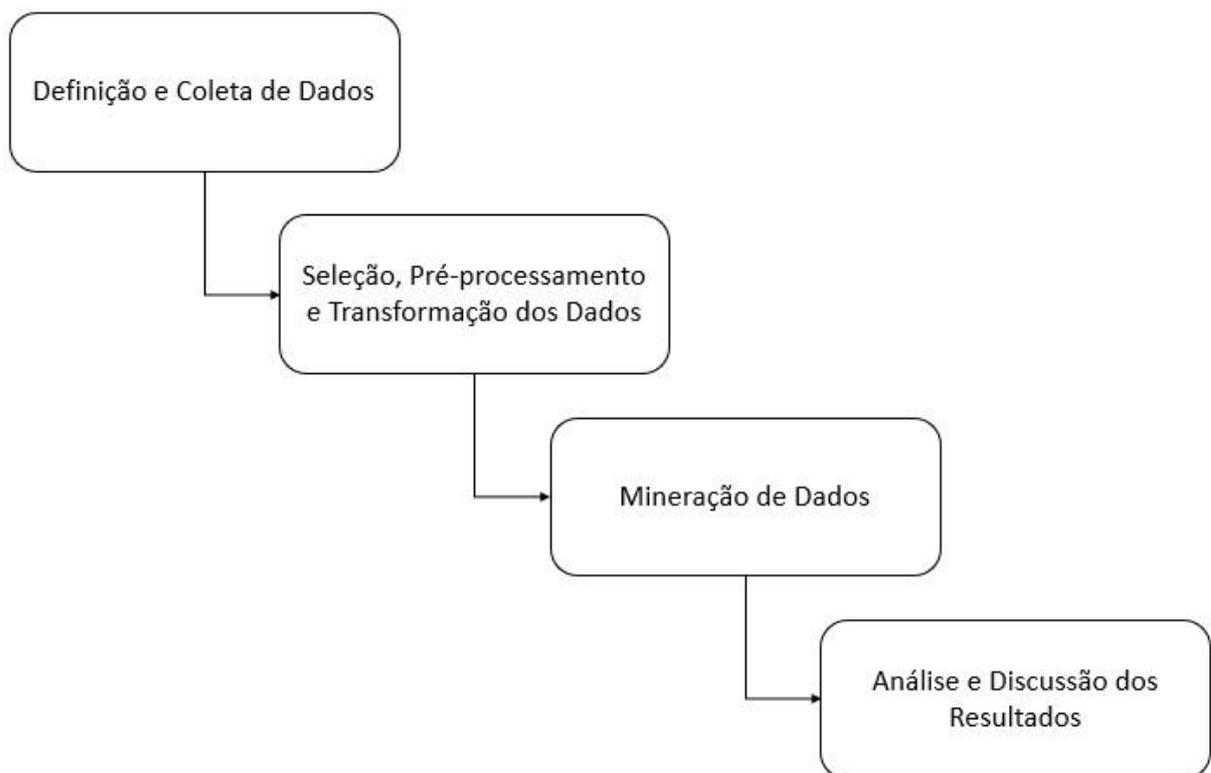
## 4.5 Metodologia

### 4.5.1 Classificação da pesquisa

A Pesquisa classifica-se do ponto de vista de sua natureza como aplicada, pois utiliza-se de técnicas consolidadas na academia para resolução de problemas e realidades locais. Do ponto de vista de sua abordagem, a pesquisa é considerada quali-quantitativa, pois é qualitativa na coleta de dados e composição das variáveis de interesse disponíveis no formulário do roteiro de fiscalização, e quantitativa no tratamento dessas variáveis, quando estas são submetidas aos algoritmos de Mineração de Dados (Silva, & Menezes, 2005).

Do ponto de vista de seus objetivos, a pesquisa é descritiva, pois a partir de um conjunto de dados, pretende-se expor a realidade presente no mesmo, mas não se intenciona explicar a sua ocorrência. Por fim, do ponto de vista de seus procedimentos, a pesquisa é classificada como levantamento na coleta de dados, bibliográfica na sua fundamentação teórica, e estudo de caso na análise do objeto de estudo (Gil, 2008).

A estratégia de investigação foi organizada e estruturada em quatro etapas sequenciais, conforme apresentado na Figura 4.1.



**Figura 4.1** Etapas da estratégia de pesquisa.

#### 4.5.2 Definição e Coleta de Dados

Os dados necessários para realização da presente pesquisa tiveram restrição de identificação, foram disponibilizados e são de propriedade do Conselho Regional de Nutricionistas da 4ª região (CRN4), e são derradeiros das fiscalizações realizadas pela equipe de fiscalização do CRN4.

O projeto de pesquisa foi submetido à Plataforma Brasil, recebendo o Certificado de Apresentação para Apreciação Ética (CAAE) sob o número 94572518.3.0000.5583 e obteve parecer consubstanciado do Comitê de Ética em Pesquisa, aprovando o projeto quanto aos aspectos éticos. Houve ainda a assinatura de Termo de Convênio de Cooperação Institucional entre o CRN4 e o autor, com o compromisso de trocar dados e informações, objetivando a melhora na fiscalização do exercício profissional, bem como das condições de trabalho, para a promoção de políticas em prol da categoria.

Todas as informações e dados técnicos exclusivos, considerados confidenciais pelo CRN4, fornecidos em função do convênio, foram recebidos e mantidos em sigilo, sendo preservado sua confidencialidade, os dados recebidos foram tratados e anonimizados, a fim de não permitir a identificação do titular das informações.

As bases de dados são constituídas a partir dos formulários dos Roteiros de Visita Técnica (RVT) disponíveis em <http://www2.crn4.org.br/pg/fiscalizacao/visitasdefiscalizacao>, que são utilizados nas fiscalizações e cadastrados no sistema próprio para tal.

As visitas técnicas são realizadas com o profissional nutricionista com a finalidade de orientação e acompanhamento do exercício profissional.

Os Roteiros de Visita Técnica são instrumentos desenvolvidos para nortear a visita do Nutricionista Fiscal, possibilitando o levantamento de informações relativas ao exercício profissional e o apoio e a orientação ao profissional. São baseados na legislação do Sistema CFN/CRN que estabelecem as atribuições dos profissionais nas diversas áreas de atuação e na legislação sanitária, buscando contemplar os principais aspectos que devem ser observados numa atuação compatível com bons padrões de qualidade (CRN4, 2019)

Para o objetivo deste trabalho, foram selecionadas as bases de dados de quatro Roteiros de Visita Técnica, a saber:

- Alimentação Coletiva - Unidade de Alimentação e Nutrição (UAN)
- Alimentação Escolar (Gestor Público)
- Nutrição Clínica - Hospital e Instituições Similares
- Nutrição Clínica - Instituição de Longa Permanência para Idosos

Os registros obtidos e utilizados, são relativos aos resultados das fiscalizações

realizadas durante o período de 2014 a 2018, abrangendo os estados do Rio de Janeiro e Espírito Santo.

O Quadro 4.1 apresenta uma breve descrição do conteúdo de alguns atributos que subdivide cada registro da base de dados coletada. São apresentados os nove primeiros atributos do arquivo original, antes das transformações.

**Quadro 4.1 - Atributos dos registros da base de dados coletada e suas respectivas descrições de conteúdo**

<b>Atributo</b>	<b>Domínio do Atributo</b>
Data Lavratura	Data + Hora da Lavratura
Fiscal	Nome do Fiscal
Tipo	Tipo de Relatório
Número do RVT	Número do relatório
Nome	Nome da empresa fiscalizada
Inscrição	Número de inscrição da empresa
CGC	Número do CGC / CNPJ da empresa
ID da Inscrição	Número de identificação do Fiscal
Vínculo de Trabalho	Tipo de vínculo de trabalho

#### **4.5.3 Seleção, Pré-processamento e Transformação dos Dados**

A etapa do pré-processamento dos dados, segundo Han, & Kamber (2011) pertence ao Processo de Extração de Conhecimento – KDD, onde a mineração de dados é uma etapa importante. Após a seleção dos dados e antes da aplicação das técnicas de mineração de dados, existe a etapa de seleção e transformação. Nesta etapa os atributos originais da base de dados são convertidos e/ou adaptados para aplicar a mineração de dados.

Nesta etapa, os dados coletados são preparados para o processo de mineração. Eles podem ser generalizados e agregados, ou até mesmo novos atributos podem ser criados ou excluídos para melhor compreensão do domínio do problema. Um processo comum é a discretização ou transformação dos dados em categorias ou intervalos quantitativos, para melhor interpretação do domínio do problema (Costa, Bernardini, & Viterbo Filho, 2014).

A base de dados disponibilizada pelo CRN4 é apresentada em arquivo Excel e vem de forma não estruturada, sendo necessário a utilização de um dicionário de dados para compatibilização, utilização e entendimento dos dados. Após a aplicação do dicionário de dados, foi feito a identificação das variáveis desfecho e não desfecho para em seguida ser gerado um arquivo de extensão CSV (Valores Separados por Vírgula, no inglês), para leitura pelo software de análise.

No fim dessa etapa, a base de dados passou a ter uma nova configuração de atributos. O

Quadro 4.2 apresenta alguns dos novos atributos, juntamente com seus respectivos domínios.

**Quadro 4.2 - Atributos e seus respectivos domínios da base de dados processada**

<b>Atributo</b>	<b>Domínio do Atributo</b>
Dia semana Lavratura	Segunda; Terça; Quarta; Quinta; Sexta; Sábado; Domingo
Mês Lavratura	1 Trimestre; 2 Trimestre; 3 Trimestre; 4 Trimestre
Hora Lavratura	Manhã; Tarde; Noite
Vínculo de Trabalho	Celetista; Concursado; Contrato; Outro
Supervisor técnico é nutricionista?	Sim; Não
Supervisor técnico	Sim; Não
RT/QT	Responsável Técnico; Quadro Técnico
Horas/dia	Até 6h; Entre 6h e 8h; Acima de 8h
Horas/semana	Até 30h; Entre 30h e 40h; Acima de 40h

O atributo “Data de Lavratura” trazia tanto a informação de data quanto a informação de hora, deste modo, foram divididos em dois atributos, um contendo a data e outro contendo a hora, entretanto, ainda assim, estes atributos acarretavam numa grande sobrecarga de informações no momento de mineração, gerando resultados que contribuíam de forma superficial para o objeto de estudo. Por isso, optou-se por substituir esses atributos, onde foi realizada a transformação discretizada, criando intervalos de valores para esses dados, onde para a data foi subdividido em dia da semana e mês de lavratura, e para o atributo de hora, foi criado intervalos de “Manhã” para horários até 12:00, “Tarde” para horários entre 12:01 e 18:00 e “Noite” para horários após 18:00. Assim, o número real que representava o dado de cada um desses atributos nos registros foi substituído pelo intervalo ao qual está inserido. Com isso, foram obtidos resultados de fácil legibilidade e compreensão, facilitando o alcance do objetivo do trabalho.

Os RVTs analisados, apresentavam ainda, uma seção com atividades obrigatórias do(a) nutricionista, subdivididos em Indicadores Qualitativos e Indicadores Quantitativos, onde cada Indicador com os atributos que os compõem, havendo ainda valores de domínio específicos para cada Indicador, a saber: “Padrão Mínimo”, “Meta Padrão” e “Não atinge padrão mínimo” para os atributos dos Indicadores Qualitativos e “Sim”, “Não”, “Às vezes” e “Não se aplica” para os atributos dos Indicadores Quantitativos. Os atributos e domínios foram transformados, consolidando-se os atributos que compõem os Indicadores em um único atributo para cada Indicador. Desta forma os novos atributos se tornaram “Indicadores Quantitativos” e “Indicadores Qualitativos” tendo como Domínios “Atende” e “Não atende”.

No Quadro 4.3 é apresentado um resumo dos domínios transformados com seus respectivos correlatos. Para a definição do desfecho final, foi considerado como “Atende” apenas os casos em que todos os atributos originais tenham sido atendidos como “Padrão mínimo”, “Meta-padrão” ou “Sim”, pois os casos onde apenas um atributo tenha sido classificado como “Não atinge padrão mínimo”, “Não” ou “Às vezes” indica que não estão sendo cumpridas as atividades obrigatórias do(a) nutricionista.

**Quadro 4.3 - Resumo dos domínios transformados com seus respectivos correlatos.**

Domínio transformado	Domínio original	
	Indicadores Qualitativos	Indicadores Quantitativos
Atende	Padrão mínimo	SIM
	Meta-padrão	
Não atende	Não atinge padrão mínimo	NÃO
		ÀS VEZES

No Quadro 4.4, é apresentada um exemplo dos atributos conforme são utilizados nas fiscalizações.

**Quadro 4.4 - Exemplo dos atributos conforme são utilizados nas fiscalizações**

INDICADORES QUALITATIVOS		*	PM	MP	NAO
4.1.	Planeja cardápios de acordo com as necessidades da clientela				
4.2.	Calcula valor nutritivo das refeições				
4.3.	Elabora/Coordena Receituário Padrão e Fichas Técnicas				
4.4.	Planeja/Executa periodicamente programas de treinamento aos colaboradores				

\* PM – Padrão mínimo MP – Meta-padrão Não – Não atinge padrão mínimo

Foram ainda transformados alguns atributos específicos de cada RVT, e que foram substituídos por intervalos levando-se em consideração às resoluções em vigor que tratam dos parâmetros numéricos mínimos de referência para atuação do nutricionista e os parâmetros numéricos mínimos no âmbito do Programa de Alimentação Escolar (PAE), deste modo, os atributos alterados e seus respectivos correlatos, foram:

- RVT Alimentação Coletiva - Unidade de Alimentação e Nutrição (UAN)

Apesar dos RVTs utilizados como base de dados estarem baseados na Resolução CFN Nº 380 (2005), que em 2018 foi revogada e atualizada para Resolução CFN Nº 600 (2018), utilizou-se como domínio dos atributos referentes à quantidade de refeições, intervalos

conforme a versão atualizada (Resolução CFN N° 600, 2018) que relacionam os parâmetros numéricos mínimos de referência para atuação do Nutricionista, conforme Quadro 4.5. Deste modo a quantidade de pequenas refeições foi convertida em grandes refeições, conforme orientado na Resolução e utilizado apenas a referência de grande refeição e os intervalos passaram a ser os mesmos utilizados como parâmetros da referida resolução.

**Quadro 4.5 - Parâmetros numéricos mínimos de referência para atuação do nutricionista - UAN.**

N° de grandes refeições/dia	Tipo de refeição		Tipo de refeição	
	Uma grande refeição/dia		Duas grandes refeições/dia ou mais	
	N° de nutricionistas	Carga horária técnica semanal	N° de nutricionistas	Carga horária técnica semanal
Até 100	1	12h	1	15h
101 a 300	1	15h	1	20h
301 a 500	1	20h	2	20h
501 a 1.000	2	30h	3	30h
1.001 a 1.500	3	30h	4	30h
1.501 de 2.500	4	30h	5	30h
Acima de 2.500	4 + 1 a cada 1.000 refeições/dia	30h	5 + 1 cada 1.000 refeições/dia	30h

**Fonte: Conselho Federal de Nutricionistas.**

A base de dados original do RVT Alimentação Coletiva – UAN, disponibilizado pelo Conselho Regional de Nutricionistas da 4ª Região para a realização deste trabalho, continham 1.734 registros, subdivididos em 164 atributos, e estes foram tratados de modo a atender os requisitos de corte temporal, bem como foram feitas as transformações pertinentes para uma melhor análise dos dados, deste modo, a base transformada passou a ter 941 registros, subdivididos em 66 atributos.

- RVT Alimentação Escolar (Gestor Público)

Os atributos referentes a quantidade de alunos também foram transformados, compatibilizando-os assim à Resolução CFN N° 465 (2010) onde relacionam os parâmetros numéricos mínimos de referência no âmbito do Programa de Alimentação Escolar (PAE) (Resolução CFN N° 465, 2010), conforme apresentado no Quadro 4.6.

**Quadro 4.6 - Parâmetros numéricos mínimos de referência no âmbito do Programa de Alimentação Escolar (PAE).**

Nº de alunos	Nº Nutricionistas	Carga horária TÉCNICA mínima semanal recomendada
Até 500	1 RT	30 horas
501 a 1.000	1 RT + 1 QT	30 horas
1.001 a 2.500	1 RT + 2 QT	30 horas
2.501 a 5.000	1 RT + 3 QT	30 horas
Acima de 5.000	1 RT + 3 QT e + 01 QT a cada fração de 2.500 alunos	30 horas

**Fonte: Conselho Federal de Nutricionistas.**

Originalmente, a base de dados referente ao RVT Alimentação Escolar (Gestor Público) possuía 164 registros, subdivididos em 172 atributos, e após as transformações e adequações, passou a apresentar 119 registros, subdivididos em 75 atributos.

- RVT Nutrição Clínica - Hospitais e Instituições Similares

Atributos referentes a número de leitos também foram transformados para adequação à resolução em vigor (Resolução CFN Nº 600, 2018), onde foram considerados as quantidades de leitos de acordo com sua complexidade, sendo os leitos de média complexidade com intervalos a cada 30 leitos e alta complexidade, a cada 15 leitos, como pode ser visto no Quadro 4.7.

**Quadro 4.7 - Parâmetros numéricos mínimos de referência para atuação do nutricionista – Hospitais e Clínicas em Geral.**

Complexidade	Nº de leitos	Nº de nutricionistas	Carga horária técnica semanal
Média	A cada 30	1	30h
Alta	A cada 15	1	30h

**Fonte: Conselho Federal de Nutricionistas.**

Para este RVT, foi disponibilizado uma base de dados com 699 registros, subdivididos em 128 atributos, que após tratado, passou a apresentar 342 registros, subdivididos em 58 atributos.

- RVT Nutrição Clínica - Instituição De Longa Permanência Para Idosos (ILPI)

As transformações realizadas referente ao RVT Nutrição Clínica - Instituição De Longa Permanência Para Idosos (ILPI) refere-se ao número de idosos atendidos, onde relaciona tanto a quantidade de nutricionista por número de idosos atendidos quanto à carga horário técnica

semanal necessária, representado no Quadro 4.8, conforme Resolução CFN Nº 600 (2018).

**Quadro 4.8 - Parâmetros numéricos mínimos de referência para atuação do nutricionista – ILPI.**

<b>Nº de idosos atendidos</b>	<b>Nº de nutricionistas</b>	<b>Carga horária técnica semanal</b>
Até 20	1	15h
De 21 a 50	1	20h
De 51 a 100	1	30h
Acima de 100	1 + 1 a cada 50 residentes	30h

**Fonte: Conselho Federal de Nutricionistas.**

Antes de sofrer as transformações, a base referente ao RVT Nutrição Clínica - Instituição De Longa Permanência Para Idosos (ILPI) contava com 214 registros, subdivididos em 176 atributos, e após passou a ter 154 registros, subdivididos em 69 atributos.

#### **4.5.4 Procedimentos metodológicos e técnicas de análises (Mineração de Dados)**

Os algoritmos foram executados a partir da ferramenta de mineração de dados WEKA (*Waikato Environment for Knowledge Analysis*), detalhes em (Frank, Hall, & Witten, 2016), ela possui algoritmos de aprendizado de máquina que podem ser utilizados para extrair informações relevantes de uma base de dados. A ferramenta foi adotada, de acordo com os seguintes motivos: i) ferramenta *open-source* e livre de custos; ii) possuir várias versões de algoritmos empregados na mineração de dados; iii) disponibilidade de recursos estatísticos para comparar o desempenho dos algoritmos e apresentar diversos recursos para análise dos dados.

Para avaliar os conhecimentos gerados com o processamento da base de estudo, a Mineração de Dados foi realizada utilizando aprendizado supervisionado para problemas de classificação, e o algoritmo classificador selecionado foi o J48, que oferece como saída um conjunto de regras de fácil interpretação por seres humanos, por meio do software WEKA. O algoritmo J48 é citado como um dos mais utilizados em mineração de dados por Wu et al. (2008) e é uma das técnicas de aprendizado de máquina que utiliza a técnica dividir-para-conquistar, onde um problema complexo é decomposto em sub-problemas mais simples, e recursivamente a mesma estratégia é aplicada a cada sub-problema (Gama, 2002).

Como resultado da aplicação do algoritmo, é apresentado o modelo encontrado pelo classificador, através da árvore de decisão, bem como sua precisão.

A utilização da árvore da decisão serve para potencializar e ajudar na tomada de

decisões de modo mais lógico e racional. O método auxilia no processo de decisões, pois permite a análise de cenários complexos e o modelo gráfico favorece a exploração de todas as possibilidades e resultados de uma determinada decisão, orientando na melhor resolução de problemas.

Na árvore de decisão, cada nó de decisão contém um teste num atributo e cada ramo descendente corresponde a um possível valor deste atributo, cada folha está associada a uma Classe e cada percurso na árvore (da raiz à folha) corresponde a uma regra de classificação.

O algoritmo J48 foi aplicado, utilizando as configurações padrões do WEKA, tendo como configurações de teste o *cross-validation* (validação cruzada) com 10 *folds* (partições de dados), de forma a avaliar a capacidade de generalização do modelo. Foram feitos ajustes nos parâmetros “confidenceFactor” e “minNumObj” com objetivo de ter uma melhor assertividade nos resultados.

No momento da mineração, para cada RVT, foram realizadas duas formas de processamento dos dados, tendo como desfecho os atributos das atividades obrigatórias do(a) nutricionista, sendo o primeiro referente aos Indicadores Qualitativos e em seguida os Indicadores Quantitativos, ou seja, em cada processamento foram geradas as árvores de decisão considerando como desfecho os atributos das atividades obrigatórias do nutricionista.

Foi ainda utilizado o Ganho de Informação, do inglês *Information Gain* para avaliar os atributos de maior relevância em relação ao desfecho escolhido, ou seja, permite criar um índice para medir qual o melhor atributo para ser o primeiro Nó da árvore, onde quanto mais alto for o valor mais robusto é o atributo em relação ao desfecho analisado.

O ganho de informação dos atributos de cada grupo de RVT serão representados em formato de quadros, separados em desfechos quantitativos e desfechos qualitativos e apresentados os 10 primeiros resultados em ordem de relevância.

Essa etapa também foi feita para cada atributo das atividades obrigatórias do(a) nutricionista, bem como para cada RVT.

Referente ao período e região estudada, esta pesquisa utilizou os RVTs registrados entre os anos de 2014 e 2018, nos estados do Rio de Janeiro e Espírito Santo, sendo assim, todos os registros que não atendiam a esses critérios, foram descartados.

#### **4.6 Análise e discussão dos resultados**

Nesta seção, são apresentados os resultados obtidos após a aplicação do algoritmo classificador J48 sobre a base de dados estudada. Os resultados foram separados em tópicos e são referentes aos seus respectivos RVTs, as análises iniciais são feitas separadamente e ao

fim é feita uma análise cruzada entre os RVTs, bem como um confronto com os trabalhos já publicados sobre a temática abordada.

#### 4.6.1 Alimentação Coletiva - Unidade de Alimentação e Nutrição (UAN)

Nesta subseção, são apresentados os resultados obtidos por meio da Mineração de Dados, realizada na base “Alimentação Coletiva UAN” entre os anos de 2014 e 2018, nos estados do Rio de Janeiro e Espírito Santo.

**Quadro 4.9 – Ganho de Informação - UAN.**

Indicadores Quantitativos		Indicadores Qualitativos	
0,061	26 Supervisiona o controle periódico de resto-ingestão	0,030	37 Relatórios técnicos / não conformidade
0,047	29 Acompanha e orienta estagiários de nutrição	0,020	29 Acompanha e orienta estagiários de nutrição
0,046	39 Avaliação qualitativa junto aos clientes	0,020	36 Procedimentos Operacionais Padronizados – POP
0,037	25 Participa do planejamento / implantação e execução de estruturas físicas da UAN	0,017	35 Manual de Boas Práticas do Serviço
0,031	37 Relatórios técnicos / não conformidade	0,015	38 Coleta de amostra das preparações
0,028	34 Programas de Informática	0,012	31 Quantidade de fontes de Atualização Técnica - Últimos 5 anos
0,022	28 Atende pacientes / clientes portadores de patologias associadas à nutrição	0,011	40 Laudo de potabilidade da água
0,022	27 Realiza estudos e pesquisas em sua área de atuação	0,011	41 Higienização da caixa d'água
0,020	35 Manual de Boas Práticas do Serviço	0,011	43 Solicitação de exames periódicos dos colaboradores
0,020	48 Layout apropriado	0,011	25 Participa do planejamento / implantação e execução de estruturas físicas da UAN

O Quadro 4.10 apresenta o Ganho de Informação, e um atributo que chama a atenção é o “Acompanha e orienta estagiários de nutrição” onde ficou bem ranqueado em ambos desfechos, sugerindo que o acompanhamento e orientação de estagiários de nutrição tem grande influência sobre o desfecho de cumprimento ou não das atividades obrigatórias do nutricionista. Tal sugestão pode ser reforçada pelo estudo realizado por Trajman et al. (2009), onde muitos profissionais reconhecem vantagens na preceptoria, tendo como pontos positivos mais mencionados a possibilidade de seu próprio desenvolvimento profissional e a melhoria do atendimento (Trajman, 2009).

Na sequência, são apresentados os resultados obtidos após a aplicação do algoritmo classificador J48 consolidado sobre a base de dados estudada, considerando tanto o Desfecho Indicadores Qualitativos quanto o Desfecho Indicadores Quantitativos.

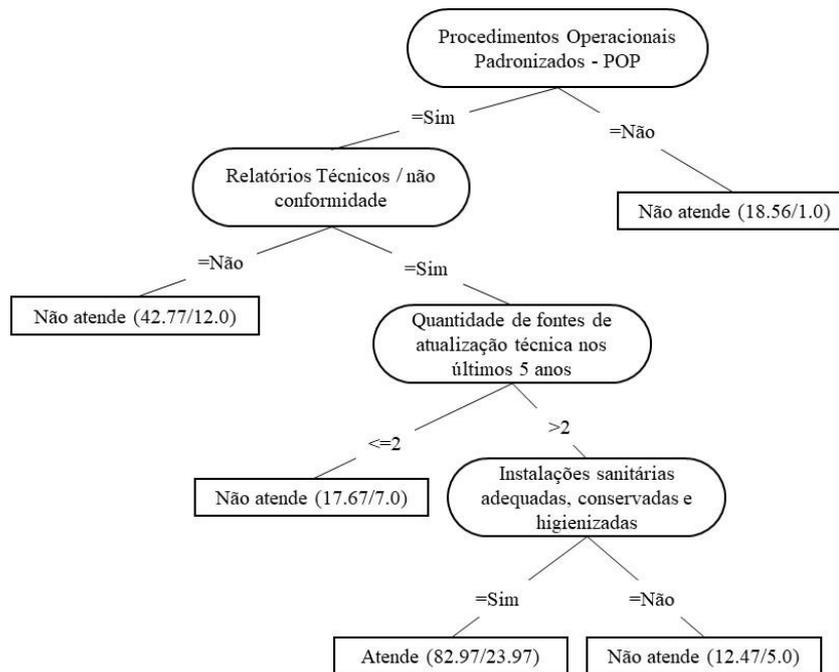
##### 4.6.1.1 Indicadores Qualitativos

Através da Árvore de Decisão é possível inferir que a utilização de Procedimentos Operacionais Padronizados (POP) é um item crítico, e a sua não utilização sugere que o

estabelecimento também não irá cumprir com as atividades obrigatórias do(a) nutricionista.

Entretanto, considerando que ocorre a utilização do POP como um recurso técnico-administrativo, a atenção passa a ser focada na utilização de Relatórios técnicos / não conformidades, que em não sendo utilizado sugere que não irá cumprir com as atividades obrigatórias do(a) nutricionista.

A quantidade de fontes de atualização técnica também se mostrou ser um fator importante de conscientização, indicando que uma maior abrangência nos canais de informação sugere uma maior taxa de cumprimento de atividades obrigatórias.



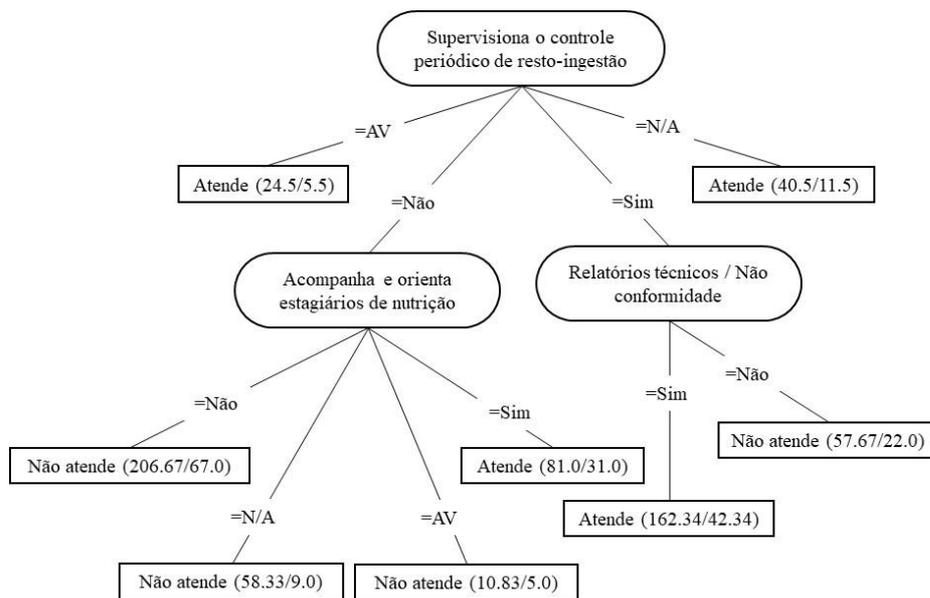
**Figura 4.2 - Árvore de Decisão - UAN - Desfecho Indicadores Qualitativos.**

#### 4.6.1.2 Indicadores Quantitativos

Para a Árvore de Decisão gerada a partir dos Indicadores Quantitativos, a indicação é de que se não há a Supervisão do controle periódico de resto-ingestão, então o Acompanhamento e orientação dos estagiários de nutrição passa a ser o fator relevante, onde os casos onde ocorrem esse acompanhamento sugerem que as atividades obrigatórias do(a) nutricionista estejam sendo cumpridas. Em contrapartida, havendo a Supervisão do controle periódico de resto-ingestão, a utilização de Relatórios técnicos / não conformidades passa a ser o direcionador, onde a sua não utilização indica o não cumprimento das atividades obrigatórias.

O indicador de controle periódico de resto-ingestão é apresentado como atividade complementar do nutricionista na UAN na Resolução CFN N° 380 (2005), porém com a atualização para a Resolução N° 600 (2018) essa atividade passou a ser obrigatória. O

resultado apresentado na árvore de decisão do controle periódico de resto-ingestão como Nó Raiz da árvore, corrobora e reforça a importância dessa atividade, como também identificado pelo próprio CFN que já passa a considerar como atividade obrigatória na nova versão da Resolução que define as áreas de atuação do nutricionista e suas atribuições, indica parâmetros numéricos mínimos de referência, por área de atuação, para a efetividade dos serviços prestados à sociedade.



**Figura 4.3 - Árvore de Decisão - UAN - Desfecho Indicadores Quantitativos.**

#### 4.6.2 Alimentação Escolar (Gestor Público)

Nesta seção, são apresentados os resultados obtidos por meio da Mineração de Dados, realizada na base “Alimentação Escolar”, inicialmente é apresentado o Ganho de informação e em seguida, as árvores de decisão para os atributos Indicadores Quantitativos e Indicadores Qualitativos.

Um atributo que chama a atenção no Ganho de Informação apresentado no Quadro 4.11 é o “Executa plano anual no âmbito do PNAE” onde ficou bem ranqueado em ambos desfechos, sugerindo que a execução do Plano Anual no âmbito do Programa Nacional de Alimentação Escolar (PNAE) tem grande influência sobre o desfecho de cumprimento ou não das atividades obrigatórias do(a) nutricionista.

**Quadro 4.10 - Ganho de Informação - Alimentação Escolar.**

Indicadores Quantitativos		Indicadores Qualitativos	
0,107	21 Executa plano anual no âmbito do PNAE	0,065	21 Executa plano anual no âmbito do PNAE
0,068	15 Participa do processo de avaliação técnica dos fornecedores de gêneros e utensílios	0,055	20 Realiza pesquisas científicas na área de nutrição
0,062	22 Titulação	0,052	15 Participa do processo de avaliação técnica dos fornecedores de gêneros e utensílios
0,055	17 Coordena ações das equipes supervisoras das unidades	0,046	9 Horas/semana
0,040	19 Acompanha e orienta estagiários de nutrição	0,043	7 RT/QT
0,030	16 Elabora fichas técnicas de produtos e preparações	0,037	2 Mês Lavratura
0,030	11 Número de Alunos	0,036	18 Identifica crianças portadoras de patologias associadas à nutrição
0,027	6 Se é supervisor técnico (sim/não)	0,033	26 Programas de informática
0,026	44 Se a visita é agendada (sim/não)	0,024	11 Número de Alunos
0,022	20 Realiza pesquisas científicas na área de nutrição	0,022	17 Coordena ações das equipes supervisoras das unidades

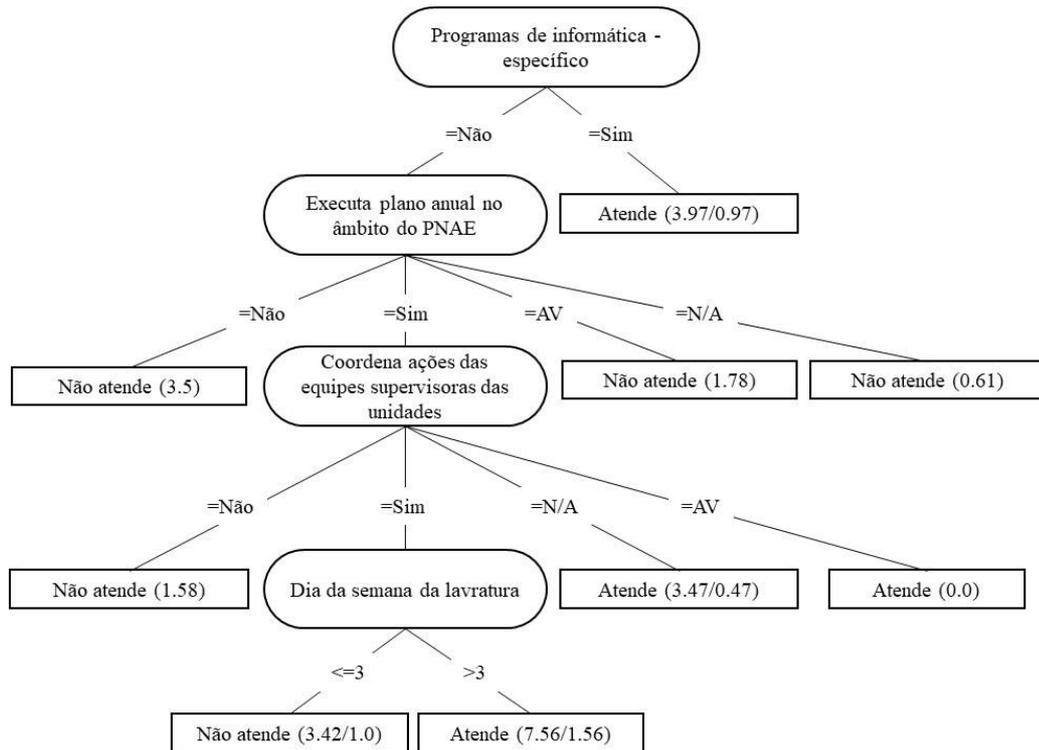
A importância deste item pode ser explicado pelo fato do plano contemplar os procedimentos adotados para o desenvolvimento das atribuições obrigatórias do(a) nutricionista no âmbito do Programa de Alimentação Escolar, que por sua vez, é pautado na Resolução Nº 465 (2010) do Conselho Federal de Nutricionistas, que dispõe sobre as atribuições do Nutricionista, estabelece parâmetros numéricos mínimos de referência no âmbito do Programa de Alimentação Escolar (PAE) e dá outras providências.

O PNAE é um programa do governo brasileiro na área de alimentação escolar e de Segurança Alimentar e Nutricional, sendo considerado um dos maiores e mais abrangentes do mundo no que se refere ao atendimento universal aos escolares e a garantia do direito humano à alimentação adequada e saudável (Ministério da Educação, 2018)

#### 4.6.2.1 Indicadores Qualitativos

A Árvore demonstra que a utilização de programas de informática específicos auxilia no cumprimento das atividades obrigatórias do(a) nutricionista, isto pois facilita e agiliza a rotina diária do nutricionista.

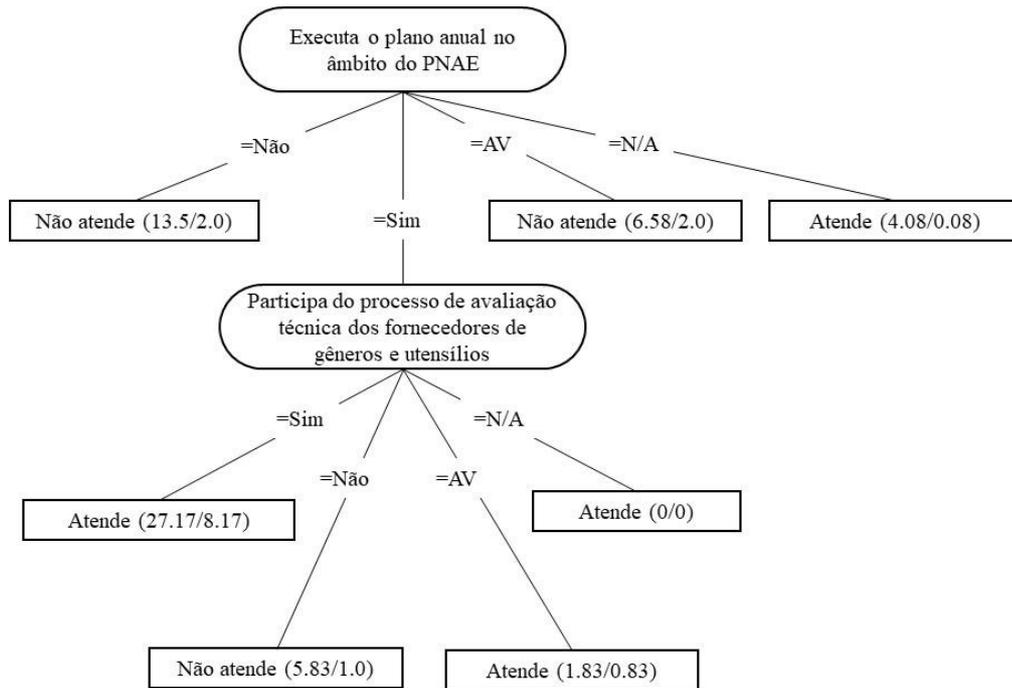
Para os casos onde não são utilizados programas específicos, a execução do plano anual no âmbito do PNAE torna-se determinante, onde apenas quando executado apresenta possibilidades de cumprimento das atividades obrigatórias.



**Figura 4.4 - Árvore de Decisão – Alimentação Escolar - Desfecho Indicadores Qualitativos**

#### 4.6.2.2 Indicadores Quantitativos

Os Indicadores Quantitativos são fortemente influenciados pela execução do plano anual no âmbito do PNAE, visto que pela árvore de decisão gerada, as situações onde a execução do plano não era realizada, apresentavam tendências em não atendimento do desfecho de realização das atividades obrigatórias, por outro lado, em caso de execução do plano, o item que passa a se tornar importante é a participação do processo de avaliação técnica dos fornecedores de gêneros e utensílios, onde caso não ocorra a participação, sugere que não se atenda também o requisito.



**Figura 4.5 - Árvore de Decisão – Alimentação Escolar - Desfecho Indicadores Quantitativos**

#### 4.6.3 Nutrição Clínica - Hospital e Instituições Similares

Nesta seção, são apresentados os resultados obtidos por meio da Mineração de Dados, realizada na base “Nutrição Clínica - Hospital e Instituições Similares”.

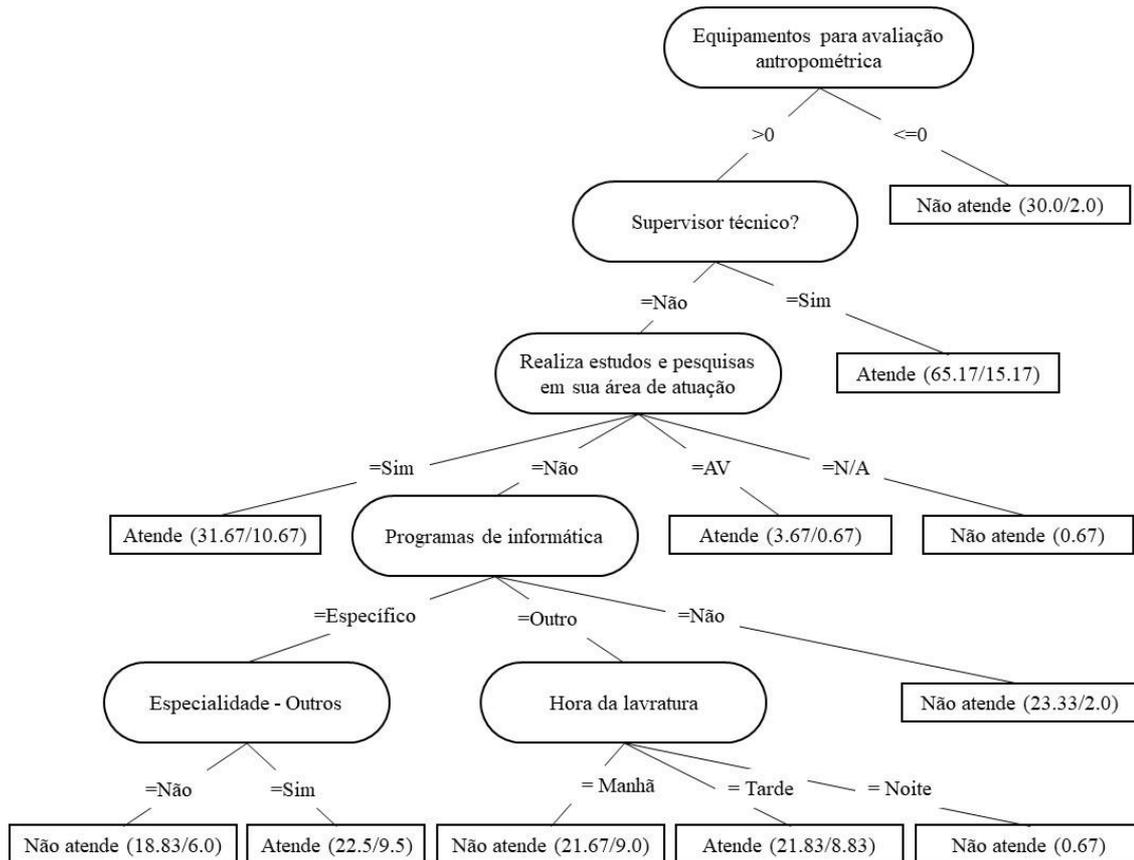
Para este RVT, não houve um único atributo que tenha se destacado consideravelmente em ambos desfechos na planilha de Ganho de Informação, porém, para os Indicadores Quantitativos, chama a atenção os três atributos mais bem ranqueados, pois tratam-se da formação acadêmica do profissional, Titulação, Especialização e as Fontes de atualização técnica, reforçando a importância na formação profissional para o cumprimento das atividades obrigatórias do(a) nutricionista. A nutrição é uma ciência, e como tal, está sempre em evolução, por isso o nutricionista deve estar sempre tecnicamente atualizado.

**Quadro 4.11 - Ganho de Informação - Hospitais e Instituições Similares.**

Indicadores Quantitativos		Indicadores Qualitativos	
0,061	38 Titulação	0,091	42 Equipamentos para Avaliação Antropométrica
0,060	40 Quantidade de Especializações	0,071	26 N Número de Leitos de média complexidade
0,058	39 Quantidade de fontes de Atualização Técnica - Últimos 5 anos	0,066	11 Se é supervisor técnico (sim/não)
0,051	26 Número de Leitos de média complexidade	0,063	41 Programas de informática
0,038	37 Acompanha e orienta estagiários de nutrição	0,052	34 Solicita exames laboratoriais necessários a avaliação nutricional
0,035	28 Número de Leitos totais	0,048	28 Número de Leitos totais
0,033	6 RT/QT	0,047	6 RT/QT
0,031	42 Equipamentos para Avaliação Antropométrica	0,045	35 Prescreve alimentos para fins especiais
0,030	41 Programas de informática	0,041	36 Realiza estudos e pesquisas em sua área de atuação
0,029	31 Horário do Plantão	0,036	39 Quantidade de fontes de Atualização Técnica - Últimos 5 anos

#### 4.6.3.1 Indicadores Qualitativos

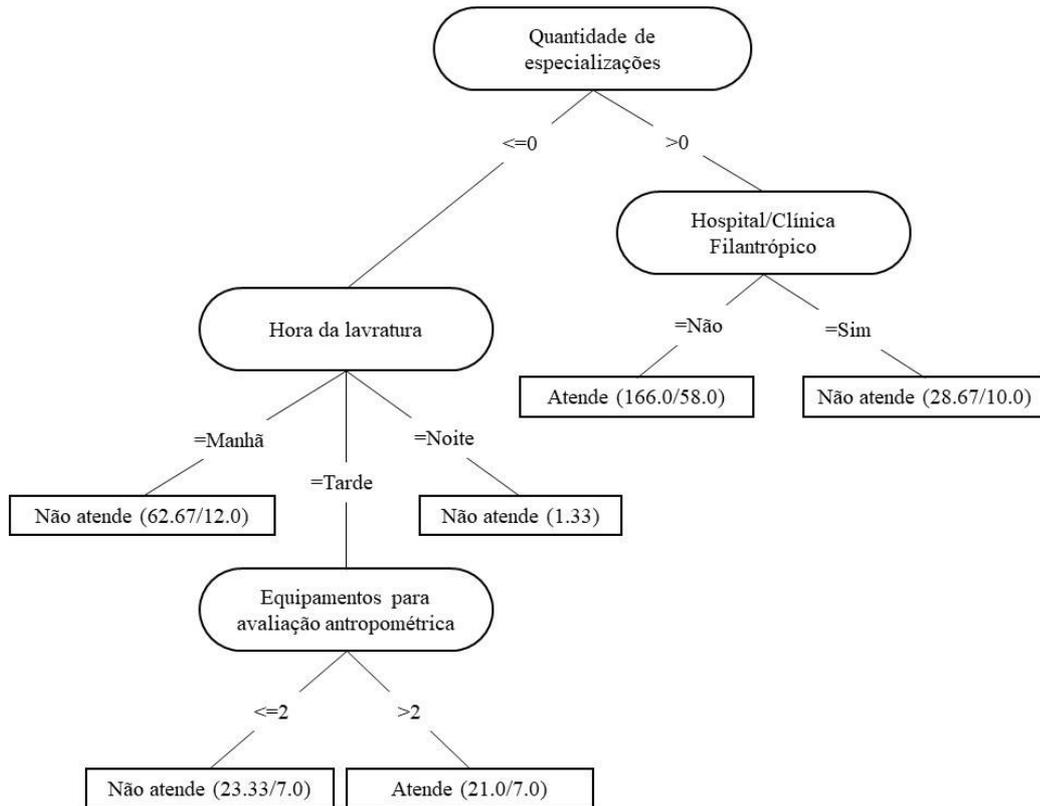
A utilização de equipamentos para avaliação antropométrica demonstrou-se importante para conclusão do desfecho qualitativo, pois quando não utilizados indicam que as atividades obrigatórias do(a) nutricionista também não estão sendo atendidos, porém, caso estejam sendo utilizados equipamentos para avaliação antropométrica, o fato de existir um supervisor técnico indica maior incidência de cumprimento das atividades obrigatórias, tal fato pode ser explicado pela presença de um superior acompanhando as atividades dos demais nutricionistas. Em não se tendo supervisor técnico, passa a ser de grande influência o fato de se realizar estudos e pesquisas em sua área de atuação, sugerindo que quando esse fato é verdadeiro, tende a se atender também as atividades obrigatórias do(a) nutricionista.



**Figura 4.6 - Árvore de Decisão – Hospitais e Instituições Similares - Desfecho Indicadores Qualitativos**

#### 4.6.3.2 Indicadores Quantitativos

Para os Indicadores Quantitativos, o item de maior relevância é a especialização do profissional, porém, caso o Hospital/Clínica seja de cunho filantrópico, a indicação sugerida é de que as atividades obrigatórias não estejam sendo cumpridas. Por outro lado, para situações com profissionais sem especialização, a hora da lavratura passou ter significância, sendo o período vespertino o mais propício a se constatar o cumprimento das atividades obrigatórias do(a) nutricionista.



**Figura 4.7 - Árvore de Decisão - Hospitais e Instituições Similares - Desfecho Indicadores Quantitativos**

#### 4.6.4 Nutrição Clínica - Instituição de Longa Permanência para Idosos (ILPI)

Nesta seção, são apresentados os resultados obtidos por meio da Mineração de Dados, realizada na base “Nutrição Clínica - Instituição de Longa Permanência para Idosos”.

O RVT Nutrição Clínica – Instituição de Longa Permanência para Idosos apresentou dois atributos relevantes no Ganho de Informação (Quadro 4.13), alternando entre primeiro e segundo lugar entre os desfechos Qualitativo e Quantitativo. Indicando a importância da prescrição de alimentos para fins especiais, bem como a avaliação qualitativa junto aos clientes.

**Quadro 4.12 - Ganho de Informação - ILPI.**

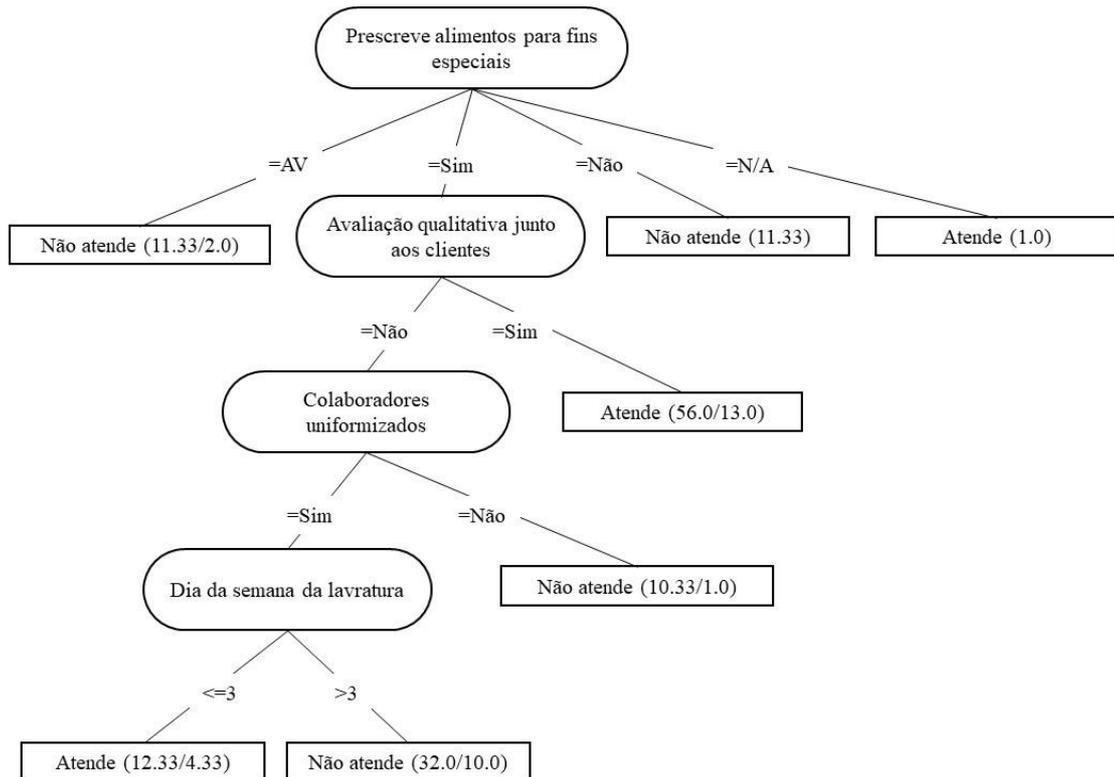
Indicadores Quantitativos		Indicadores Qualitativos	
0,085	24 Avaliação qualitativa junto aos clientes	0,112	20 Prescreve alimentos para fins especiais
0,057	20 Prescreve alimentos para fins especiais	0,091	24 Avaliação qualitativa junto aos clientes
0,056	14 Local visitado	0,058	14 Local visitado
0,052	31 Lavatórios e produtos de assepsia para manipuladores	0,040	25 Laudo de potabilidade da água
0,049	32 Controle de temperatura - impresso	0,032	46 Equipamentos para Avaliação Antropométrica
0,048	8 Horas/semana	0,029	26 Higienização da caixa d'água
0,042	30 Desinfecção de gêneros	0,026	18 Realiza estudos e pesquisas em sua área de atuação
0,039	28 Solicitação de exames periódicos dos colaboradores	0,023	19 Solicita exames laboratoriais necessários a avaliação nutricional
0,037	44 Supervisiona o controle periódico de resto-ingestão	0,021	43 Participa do planejamento /implantação e execução de estrutura física da UAN
0,029	48 Manual de Boas Práticas	0,020	29 Colaboradores uniformizados

#### 4.6.4.1 Indicadores Qualitativos

A árvore de decisão vista na Figura 4.8 indicou como atributo mais importante a prescrição de alimentos para fins especiais, sugerindo que a não prescrição ou a inconsistência na prescrição acarreta o não cumprimento das atividades obrigatórias do(a) nutricionista. Sequencialmente o próximo atributo a se observar é referente à avaliação qualitativa junto aos clientes, sugerindo que a execução de pesquisas de satisfação reflete no cumprimento das atividades obrigatórias.

A prescrição de alimentos para fins especiais na faixa etária senil é um item importante pela diminuição natural e progressiva de massa muscular presente na referida faixa etária, conhecido como Sarcopenia, indicando que o cumprimento das atividades obrigatórias irá fornecer informações que irão propiciar a identificação e triagem de risco nutricional.

Um estudo realizado por Landi et al. (2011) indica a alta prevalência de sarcopenia entre os residentes de Instituições de Longa Permanência para Idosos, identificando o risco nutricional e a diminuição da massa muscular nesses idosos e relata a importância da suplementação nutricional nessa faixa etária para correção e prevenção da sarcopenia. O resultado do estudo feito por Landi et al. (2011) confirma a importância ressaltada pela árvore de decisão, onde indicam a prescrição de alimentos para fins especiais como atributo de maior relevância, visto que, como também relatado por Patronillo et al. (2015) a suplementação proteica deve ser considerada nos idosos em risco nutricional, ou com sarcopenia já estabelecida, que não atinjam a ingestão proteica diária, necessária para suprir as necessidades metabólicas, através de modificações na dieta, e nos quais não existam contraindicações médicas.



**Figura 4.8 - Árvore de Decisão - ILPI - Desfecho Indicadores Qualitativos**

#### 4.6.4.2 Indicadores Quantitativos

O desfecho Quantitativo de Nutrição Clínica em ILPI tem como principal atributo a avaliação qualitativa junto aos clientes, onde o simples fato de se executar uma pesquisa de satisfação junto aos clientes torna-se um indicativo de cumprimento das atividades obrigatórias do(a) nutricionista, porém, a não realização da avaliação ressalta como importante o indicativo de utilização ou não de uniformes pelos colaboradores, onde a não utilização sugere que também não estejam sendo cumpridas as atividades obrigatórias. No caso de utilização de uniformes, o próximo atributo a se tornar relevante para o desfecho Quantitativo passa a ser a existência de lavatórios e produtos de assepsia para manipuladores, ressaltando que o cumprimento desse atributo indica maior probabilidade de cumprimento das atividades obrigatórias do(a) nutricionista.

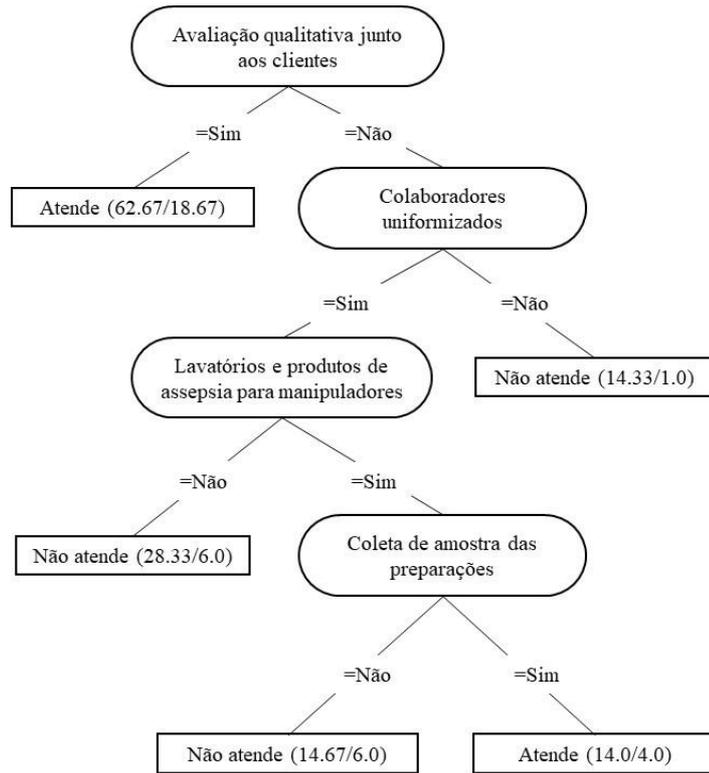


Figura 4.9 - Árvore de Decisão - ILPI - Desfecho Indicadores Quantitativos

#### 4.6.5 Comparativo entre RVTs

Na figura 4.10 é apresentada a distribuição temporal da quantidade de RVT entre os anos de 2014 e 2018, nos estados do Rio de Janeiro e Espírito Santo.



Figura 4.10 - Quantidade de RVTs por ano.

Legenda:

UAN = RVT Alimentação Coletiva – UAN

AE = RVT Alimentação Escolar (Gestor Público)

**HES = RVT Nutrição Clínica - Hospitais e Instituições Similares**

**ILPI = RVT Nutrição Clínica - Instituição de Longa Permanência Para Idosos**

A distribuição da quantidade de RVTs ao longo do tempo se concentra mais entre os anos de 2016 e 2017, onde em 2016 teve um significativo aumento no número de RVTs, crescendo um pouco mais em 2017, porém recuando novamente em 2018, mas ainda em patamar superior aos anos de 2014 e 2015.

O gráfico mostra também uma diferença considerável na quantidade de RVTs relativos à Alimentação Coletiva – UAN em relação aos demais, sendo superior ao dobro do somatório dos outros três.

No Quadro 4.11 são apresentadas as precisões específicas de cada RVT considerando os atributos “Indicador Quantitativo”, abreviado no Quadro como “Ind. Quant.” e “Indicador Qualitativo”, abreviado no Quadro como “Ind. Qual.”.

**Quadro 4.13 - Valores de Precisão em cada desfecho.**

	UAN		ALIMENTAÇÃO ESCOLAR		HOSPITAIS E SIMILARES		ILPI	
	Ind. Quant.	Ind. Qual.	Ind. Quant.	Ind. Qual.	Ind. Quant.	Ind. Qual.	Ind. Quant.	Ind. Qual.
Não atende	0,765	0,950	0,808	0,925	0,667	0,747	0,676	0,720
Atende	0,507	0,204	0,348	0,179	0,587	0,506	0,566	0,667
Média ponderada	0,673	0,868	0,692	0,844	0,631	0,658	0,626	0,695

Através dos resultados apresentados no Quadro 4.14, pode ser observado que a precisão para a classificação “Não atende” foi superior a precisão para a classificação “Atende” em todos os atributos de todos os RVTs, e isso é um fator positivo para a aplicação e usabilidade do modelo decisório apresentado neste artigo, visto que, por se tratarem de ações fiscais, torna-se mais relevante a precisão para a definição da não conformidade, particularmente a classificação “Não atende” do que a precisão para a definição da conformidade, no caso a classificação “Atende”.

Este trabalho apresentou uma forma de predição do cumprimento de atividades obrigatórias, utilizando para isso dados de fiscalizações anteriores, que são observações verificadas no local visitado mediante Roteiro de Visita Técnica e efetuado por profissional capacitado para a fiscalização, deste modo auxiliando e podendo servir de direcionamento para futuras fiscalizações.

Não foi encontrado na literatura trabalhos que utilizem mineração de dados da

fiscalização formal para a verificação do exercício profissional, o que foi encontrado são trabalhos que utilizam a mineração com as avaliações de comentários em texto aberto, feito por profissionais pares como faz Gibbons et al. (2017); ou o banco de dados resultante do exercício profissional como apresentado por Hautemanière et al. (2011) que utiliza um banco de dados de mortalidade de um hospital universitário, a fim de verificação da conduta profissional.

Outros autores utilizam estatística para comparar resultados de modelo estatístico versus avaliação clínica feita por especialistas, como faz Bernstein et al. (2017) em seu trabalho que utiliza dados dos pacientes para fazer as comparações. Layzell (2005) contextualiza sobre prescrição inadequada de medicamentos em pacientes pós-operatório, com o objetivo de identificar, através de auditoria feita durante 5 semanas em uma enfermaria de recuperação, os atrasos sofridos pelos pacientes para o alívio da dor no pós-operatório e os motivos.

Esses trabalhos mostram avaliações não direcionadas e algumas até pontuais, onde não há roteiro ou padrão de verificação ou continuidade, podendo apresentar resultados enviesados ou superficiais. Outrora, nenhum deles busca dar subsídios para órgãos oficiais de fiscalização.

O trabalho que mais se aproxima desse objetivo é o estudo feito por Wang, Balasubramani, & Cruz (2017) que utiliza avaliações on-line de consumidores de restaurantes dos Estados Unidos e introduzem a visão de construir um sistema de inspeção em tempo real combinado com inspeções de rotina do departamento de saúde pública para impedir que os usuários sofram doenças transmitidas por alimentos. Os autores exploram análises com conjuntos de palavras-chave que denotam doenças transmitidas por alimentos e utilizam técnicas de aprendizado supervisionado para classificar os dados, a fim de prever se uma avaliação indica doenças transmitidas por alimentos. Em particular, detectam avaliações que preveem doenças transmitidas por alimentos por recursos e métodos de classificação selecionados.

#### **4.7 Considerações Finais**

Esta pesquisa se propôs a analisar o comportamento dos dados resultantes das ações fiscais do Conselho Regional de Nutricionistas da 4ª Região. A aplicação do algoritmo classificador J48, retornou resultados satisfatórios, com precisão de classificação para a não conformidade bem superiores do que a precisão de classificação para a conformidade, principalmente para os Indicadores Qualitativos, onde apresentou até 95% de acurácia para a classificação “Não atende”, o que expressa a confiabilidade das informações geradas.

Os altos valores de porcentagens alcançados nos parâmetros de acurácia indicam que o nível de confiabilidade dos resultados obtidos é satisfatório, o que permite a realização de inferências que condizem com a realidade do objeto de estudo.

De acordo com os resultados obtidos, podemos concluir que é possível encontrar uma tendência de resultado e possíveis padrões nos desfechos propostos para o processo de mineração de dados utilizando-se a base de dados oriunda de fiscalizações anteriores.

Para cada RVT tendem a ressaltar como mais importantes e influentes atributos relativos e específicos da área de atuação em que o próprio RVT está inserido, como por exemplo a execução do plano anual no âmbito do PNAE para a área de Alimentação Escolar e a prescrição de alimentos para fins especiais para Instituições de Longa Permanência, indicando a particularidade de cada área.

Esse estudo pode servir de material para ações de fiscalizações futuras, direcionando e auxiliando na priorização dos itens de verificação no momento das visitas de fiscalização, auxiliando na identificação de irregularidades e não conformidades.

#### 4.8 Referências

- Barbosa, D. C. C., & Machado, M. A. (2007). Mineração de Dados usando o software WizRule em Base de Dados de Compras de TI. *Revista Eletrônica de Sistemas de Informação*, 6(1). <https://doi.org/10.21529/RESI.2007.0601001>
- Bernstein, D.N., Keswani, A., Ring, D. (2017). Perioperative Risk Adjustment for Total Shoulder Arthroplasty: Are Simple Clinically Driven Models Sufficient?. *Clinical Orthopaedics and Related Research*. <https://doi.org/10.1007/s11999-016-5147-y>
- Berry, M. J. A., & Linoff, G. (2004). Data mining techniques: For marketing, sales, and customer relationship management (2nd ed). Indianapolis, Ind: Wiley Pub.
- Berson, A., Smith, S., & Thearling, K. (2000). Building data mining applications for CRM. New York, NY: McGraw-Hill.
- Camilo, C. O., & Silva, J. C. da. (2009). Mineração de Dados: Conceitos, Tarefas, Métodos e Ferramentas. 29.
- Costa, J. D. J., Bernardini, F. C., Viterbo Filho, J. (2014). A mineração de dados e a qualidade de conhecimentos extraídos dos boletins de ocorrência das rodovias federais brasileiras. *AtoZ: novas práticas em informação e conhecimento*, v.3, n.2, p.139, 31 dez 2014.
- CRN4 (2019). Fiscalização. Visitas de Fiscalização. Disponível em: <<http://www2.crn4.org.br/pg/fiscalizacao/visitasdefiscalizacao>>. Acessado em 08/12/2019.
- Frank E., Hall M. A., & Witten I. H. (2016). The WEKA Workbench. Online Appendix for “Data Mining: Practical Machine Learning Tools and Techniques”. *Morgan Kaufmann*.
- Gama, J. (2002). Árvores de Decisão. Disponível em: <[https://www.dcc.fc.up.pt/~ines/aulas/MIM/arvores\\_de\\_decisao.pdf](https://www.dcc.fc.up.pt/~ines/aulas/MIM/arvores_de_decisao.pdf)>. Acessado em: 01/02/2020.
- Gibbons, C., Richards, S., Valderas, J.M., Campbell, J. (2017). Supervised machine learning algorithms can classify open-text feedback of doctor performance with human-level accuracy. *Journal of Medical Internet Research*. <https://doi.org/10.2196/jmir.6533>
- Gil, A. C. (2008). Métodos e técnicas de pesquisa social. São Paulo: Atlas.
- Han, J., & Kamber, M. (2011). Data mining: Concepts and techniques (3rd ed). Burlington, MA: Elsevier.
- Hautemanière, A., Florentin, A., Hartemann, P., & Hunter, P.R. (2011). Identifying possible deaths associated with nosocomial infection in a hospital by data mining. *American Journal of Infection Control*. <https://doi.org/10.1016/j.ajic.2010.04.216>
- Landi F., Liperoti R., Fusco D., Mastropaolo S., Quattrociochi D., Proia A., Russo A., Bernabei R., & Onder G. (2012). Prevalence and Risk Factors of Sarcopenia Among

- Nursing Home Older Residents, *The Journals of Gerontology: Series A*, Volume 67A, Issue 1, January 2012, Pages 48–55, <https://doi.org/10.1093/gerona/glr035>
- Layzell, M. (2005). Improving the management of postoperative pain. *Nursing times*. 2005 Jun 28-Jul 4;101(26):34-6.
- Lei Nº 6.583. Cria os Conselhos Federal e Regionais de Nutricionistas, regula o seu funcionamento, e dá outras providências. Pub. L. No. LEI No 6.583 (1978).
- Lobo, L. C. (2017). Inteligência Artificial e Medicina. *Revista Brasileira de Educação Médica*, 41(2), 185–193. <https://doi.org/10.1590/1981-52712015v41n2esp>
- Ministério da Educação (2018). Manual de apoio para as atividades técnicas do Nutricionista no âmbito do PNAE. *Fundo Nacional de Desenvolvimento da Educação*. Brasília, DF.
- Neves, B., Raimundo, A., & Obermeyer, Z. (2017). The Quiet Revolution of Big Data in Medicine. *Medicina Interna*, 24(4). <https://doi.org/10.24950/rspmi/Perspective/2017>
- Patronillo C., Veríssimo M. T. (2015). Suplementos proteicos e Sarcopenia no idoso – Artigo de Revisão. *Faculdade de Medicina*. Universidade de Coimbra, Portugal.
- Pina, M. de F., & Carvalho, M. S. (2017). GeoMed 2017: Visão mais profunda a partir de big data e pequenas áreas. *Cadernos de Saúde Pública*, 33(10). <https://doi.org/10.1590/0102-311x00172017>
- Resolução CFN Nº 380 (2005). Dispõe Sobre a Definição das Áreas de Atuação do Nutricionista e Suas Atribuições, Estabelece Parâmetros Numéricos de Referência, Por Área de Atuação, e dá Outras Providências. *Conselho Federal de Nutricionistas*. Pub. L. No. RESOLUÇÃO CFN No 380.
- Resolução CFN Nº 465 (2010). Dispõe sobre as atribuições do Nutricionista, estabelece parâmetros numéricos mínimos de referência no âmbito do Programa de Alimentação Escolar (PAE) e dá outras providências. *Conselho Federal de Nutricionistas*. Pub. L. No. RESOLUÇÃO CFN No 465 (2010).
- Resolução CFN Nº 600 (2018). Dispõe sobre a definição das áreas de atuação do nutricionista e suas atribuições, indica parâmetros numéricos mínimos de referência, por área de atuação, para a efetividade dos serviços prestados à sociedade e dá outras providências. *Conselho Federal de Nutricionistas*. Pub. L. No. RESOLUÇÃO CFN No 600 (2018).
- Silva, E. L. da, & Menezes, E. M. (2005). Metodologia da Pesquisa e Elaboração de Dissertação. (4. ed. rev. atual.), 138.
- Trajman, A., Assunção, N., Venturi, M., Tobias, D., Tochi, V., & Brant, V (2009). A preceptoria na rede básica da Secretaria Municipal de Saúde do Rio de Janeiro: opinião dos profissionais de saúde. *Revista Brasileira de Educação Médica*. Rio de Janeiro, v.33, n. 1, p. 24-31, 2009.
- Wang, Z., Balasubramani, B. S., & Cruz, I. F. (2017). Predictive Analytics Using Text Classification for Restaurant Inspections. In *ACM SIGSPATIAL Workshop on Smart Cities and Urban Analytics (UrbanGIS)*. Article No.: 14 Pages 1–4. <https://doi.org/10.1145/3152178.3152192>

Weiss, S. M., & Indurkha, N. (1998). *Predictive data mining: A practical guide*. San Francisco: *Morgan Kaufmann Publishers*.

Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., & Steinberg, D. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, 14(1), 1–37. <https://doi.org/10.1007/s10115-007-0114-2>

## 5 CONSIDERAÇÕES FINAIS

Este trabalho iniciou-se com uma pesquisa bibliométrica sobre estudos que abordam os temas de mineração de dados, fiscalização e exercício profissional, em que foi identificado que esta é uma área em ascensão, pertinente e com muito a explorar. Para a definição da metodologia, foi analisado um conjunto de trabalhos, onde a partir daí foi possível analisar que os trabalhos na área da saúde ressaltaram como os mais frequentes.

Por meio da análise sistemática, após a inclusão dos termos relativos à área da saúde, foi verificada, apesar de existirem poucos trabalhos na área, a viabilidade de utilização de ferramentas matemáticas como forma de verificação ou apoio do exercício profissional. Buscou-se a análise e discussão dos resultados com a apresentação das ferramentas matemáticas utilizadas pelos autores.

Na realização do estudo de caso no Conselho Regional de Nutricionistas da 4ª Região, foram analisados os bancos de dados disponibilizados pelo próprio CRN4, provindos de visitas de fiscalização realizadas. A principal conclusão é que, é possível encontrar uma tendência de resultado e possíveis padrões nos desfechos propostos para o processo de mineração de dados e, em geral, os atributos mais relevantes são os relativos e específicos da área de atuação em que o próprio RVT está inserido.

Tendo em vista a carência de estudos específicos envolvendo técnicas de mineração e fiscalização do exercício profissional, este trabalho contribui, apresentando a viabilidade de extração de informações estratégicas, a partir da própria base de dados do Conselho Profissional. Acredita-se que os resultados obtidos sejam de interesse do CRN4 e também de outros Conselhos Profissionais, de modo que o método aqui apresentado pode ser reutilizado em outras bases de dados e servir de material para ações de fiscalizações futuras, direcionando e auxiliando na priorização dos itens de verificação no momento das visitas de fiscalização, auxiliando na identificação de irregularidades e não conformidades.

## REFERÊNCIAS GERAIS

- Carvalho, D. R., Moser, A. D., Silva, V. A. da, & Dallagassa, M. R. (2012). Mineração de Dados aplicada à fisioterapia. *Fisioterapia em Movimento*, 25(3), 595–605. <https://doi.org/10.1590/S0103-51502012000300015>
- Rezende, S. O. (Org.). (2003). *Sistemas inteligentes: Fundamentos e aplicações* (1. ed). Barueri, SP: Ed. Manole.