

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO**

Daniel Antonio da Silva

**MINERAÇÃO DE TEXTOS APLICADA NA PREVISÃO E
DETECÇÃO DE EVENTOS ADVERSOS NO HOSPITAL
DE CLÍNICAS DE PORTO ALEGRE**

Porto Alegre

2017

Daniel Antonio da Silva

**Mineração de textos aplicada na previsão e detecção de eventos adversos no
Hospital de Clínicas de Porto Alegre**

Dissertação submetida ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal do Rio Grande do Sul como requisito parcial à obtenção do título de Mestre em Engenharia de Produção, modalidade Acadêmica, na área de concentração em Sistemas de Qualidade.

Orientadora: Prof^a. Carla Schwengber ten Caten, Dra.

Porto Alegre

2017

**Mineração de textos aplicada na previsão e detecção de eventos adversos no
Hospital de Clínicas de Porto Alegre**

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Engenharia de Produção na modalidade Acadêmica e aprovada em sua forma final pelo Orientador e pela Banca Examinadora designada pelo Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal do Rio Grande do Sul.

Prof^a. Carla Schwengber ten Caten, Dra.
Orientador PPGEP/UFRGS

Prof. Flávio Sanson Fogliatto
Coordenador PPGEP/UFRGS

Banca Examinadora:

Prof. Flávio Sanson Fogliatto, *Ph.D.* (PPGEP/UFRGS)

Prof. José Miguel Dora, Dr. (HCPA)

Prof. Rodrigo Pires dos Santos, Dr. (HCPA)

Dedico esta dissertação
à minha família e amigos.

AGRADECIMENTOS

Agradeço à professora Carla por todo o apoio, conhecimento transmitido e por ter incentivado esta pesquisa desde o início.

Agradeço à CAPES pelo apoio financeiro durante todo o período deste mestrado.

Agradeço à equipe do HCPA, Rodrigo Santos, Márcia Pires e Tiago Vaz, sem os quais esta pesquisa não seria possível. Agradeço também a Dra. Helena por direcionar e incentivar este trabalho desde o início.

Agradeço aos colegas e professores da UFRGS, PPGE e LOPP e também ao colega Filipe Lucini pelas contribuições a este estudo.

Agradeço à minha família, especialmente à minha mãe Marlene, pela educação, apoio incondicional e amor transmitidos durante toda a minha vida.

Agradeço à Sílvia, Joana e Thomas pelo apoio e hospitalidade nesta etapa final.

Agradeço especialmente à minha noiva Tamar por sempre acreditar, me apoiar nos meus projetos e estar sempre ao meu lado. Agradeço por me fazer acreditar que podemos tornar o mundo um *lugar* melhor.

RESUMO

Este trabalho apresenta os resultados de uma pesquisa que teve como objetivo avaliar o desempenho de métodos de mineração de textos na previsão e detecção de Eventos Adversos (EA). A primeira etapa foi a revisão sistemática da literatura que buscou identificar os métodos de mineração de textos e as áreas da saúde que esses estão sendo aplicados para prever e detectar EA. Após essa etapa foi realizada uma aplicação de métodos de mineração de textos para prever Infecções do Sítio Cirúrgico (ISC) a partir do texto livre de descrições cirúrgicas no Hospital de Clínicas de Porto Alegre (HCPA). Por fim, métodos de mineração de textos foram aplicados para detectar ISC a partir do texto das evoluções de pacientes 30 (trinta) dias após uma cirurgia. Como resultados, destaca-se a identificação dos melhores métodos de pré-processamento e mineração de textos para prever e detectar ISC no HCPA, podendo ser aplicados a outros EA. O método Stochastic Gradient Descent (SGD) apresentou o melhor desempenho, 79,7% de ROC-AUC na previsão de EA. Já para detecção de EA o melhor método foi o Logistic Regression, com desempenho 80,6% de ROC-AUC. Os métodos de mineração de textos podem ser usados para apoiar de maneira eficaz a previsão e detecção de EA, direcionando ações de vigilância para a melhoria da segurança do paciente.

Palavras-chave: Eventos Adversos; Infecção do Sítio Cirúrgico; Mineração de textos.

ABSTRACT

This work presents the results of a research that aimed to evaluate the performance of text mining methods in the prediction and detection of Adverse Events (AE). The first step was the systematic review of the literature that sought to identify the methods of text mining and the health areas they are being applied to predict and detect AE. After this step, an application of text mining methods was performed to predict Surgical Site Infections (SSI) from the free text of medical records at Hospital de Clínicas de Porto Alegre (HCPA). Finally, text mining methods were applied to detect SSI from the text of medical records 30 (thirty) days after surgery. As results, is highlight the identification of the best methods of pre-processing and text mining to predict and detect SSI in the HCPA, and can be applied to other AE. The Stochastic Gradient Descent (SGD) presented the best performance, 79.7% of ROC-AUC in the prediction of AE. Already for the detection of AE the best method was the Logistic Regression, with performance 80.6% of ROC-AUC. Text mining methods can be used to effectively support the prediction and detection of AE by directing surveillance actions to improve patient safety.

Key words: Adverse Events. Surgical Infection. Text Mining.

LISTA DE FIGURAS

Figura 1. Processo cirúrgico.	16
Figura 2. Publicações selecionadas	25
Figura 3. Procedimentos Metodológicos.....	59
Figura 4. Matriz de termos.....	63
Figura 5. ROC-AUC: AD: AdaBoost, DT: DecisionTree, LSVC: LinearSVC, LR: Logistic Regression, NB: MultinomialNB, NC: Nearest Centroid, RF: Random Forest, SGD: Stochastic Gradient Descent, SVC: Support Vector Classification	70
Figura 6. Redução do volume de dados vs. Falsos Negativos.	71
Figura 7. Cirurgias infectadas existentes e previstas.....	71
Figura 8. Procedimentos Metodológicos.....	82
Figura 9. ROC-AUC. LSVC: LinearSVC, LR: Logistic Regression, NB: MultinomialNB, NC: Nearest Centroid, RF: Random Forest, SGD: Stochastic Gradient Descent, SVC: Support Vector Classification.	89
Figura 10. Redução do volume de dados vs. Falsos Negativos.	90

LISTA DE TABELAS

Tabela 1. Questões de Pesquisa.....	26
Tabela 2. Registros textuais selecionados.....	68
Tabela 3. Desempenho dos classificadores. SA: Seleção de Atributos, P: Percentil, T: Transformação, N: Normalização, CW: Peso / Probabilidade Prévia, DP: Desvio Padrão.	69
Tabela 4. Registros textuais selecionados.....	88
Tabela 5. Desempenho dos classificadores. SA: Seleção de Atributos, P: Percentil, T: Transformação, N: Normalização, CW: Peso / Probabilidade Prévia, DP: Desvio Padrão.	89

LISTA DE SIGLAS E ABREVIATURAS

AM	Aprendizagem de Máquina
ANVISA	Agência Nacional de Vigilância Sanitária
ASA	American Society of Anesthesiologists
CCIH	Comissão de Controle de Infecção Hospitalar
DCBD	Descoberta de Conhecimento em Bases de Dados
EA	Eventos Adversos
EM	Erro Médico
FAERS	FDA Adverse Event Report System
FDA	Food and Drug Administration
FN	Falsos Negativos
FP	Falsos Positivos
HCPA	Hospital de Clínicas de Porto Alegre
IDF	Frequência Inversa de Documentos
IHI	Institute for Healthcare Improvement
IOM	Institute of Medicine
IRIC	Índice de Risco de Infecção Cirúrgica
ISC	Infecções de Sítio Cirúrgico
LinearSVC	Linear Support Vector Classification
MT	Mineração de Textos
NLTK	Natural Language Toolkit
OMS	Organização Mundial da Saúde
PLN	Processamento de Linguagem Natural
PPC	Portal de Periódicos da Capes
RAM	Reações Adversas a Medicamentos
REH	Registros Eletrônicos Hospitalares
REN	Reconhecimento de Entidades Nomeadas
ROC-AUC	Área Sob a Curva das Características Operacionais do Receptor
SGD	Stochastic Gradient Descent
SVC	Support Vector Classification
TF	Frequência do Termo
TN	Verdadeiros Negativos
TP	Verdadeiros Positivos
VAERS	Vaccine Adverse Event Reporting System

SUMÁRIO

1	INTRODUÇÃO.....	13
1.1	TEMA E OBJETIVOS.....	14
1.2	JUSTIFICATIVA DO TEMA E OBJETIVOS.....	15
1.3	DELINEAMENTO DO ESTUDO	16
1.3.1	MÉTODO DE PESQUISA.....	17
1.3.2	MÉTODO DE TRABALHO	17
1.4	DELIMITAÇÕES DO ESTUDO	18
1.5	ESTRUTURA DA DISSERTAÇÃO.....	18
1.6	REFERÊNCIAS	20
2	ARTIGO 1: Detecção de eventos adversos por meio da mineração de textos: uma revisão sistemática da literatura.....	22
2.1	INTRODUÇÃO.....	22
2.2	PROCEDIMENTOS METODOLÓGICOS	24
2.2.1	QUESTÕES DE PESQUISA	25
2.3	REVISÃO SISTEMÁTICA DA LITERATURA.....	26
2.3.1	TRABALHOS CORRELATOS	26
2.3.2	MINERAÇÃO DE TEXTOS	27
2.3.3	DETECÇÃO DE EVENTOS ADVERSOS.....	29
2.3.4	FARMACOVIGILÂNCIA	33
2.3.4.1	FARMACOVIGILÂNCIA EM REDES SOCIAIS	33
2.3.4.2	REAÇÕES ADVERSAS A MEDICAMENTOS	34
2.3.4.3	REAÇÕES ADVERSAS A VACINAS.....	37
2.3.4.4	INTERAÇÕES MEDICAMENTOSAS	38
2.4	CONSIDERAÇÕES FINAIS	38
2.5	REFERÊNCIAS	41
2.6	REFERÊNCIAS DA REVISÃO SISTEMÁTICA.....	44
3	ARTIGO 2: Previsão de Infecções do Sítio Cirúrgico por meio da mineração de textos no Hospital de Clínicas de Porto Alegre.....	54
3.1	INTRODUÇÃO.....	54
3.2	REFERENCIAL TEÓRICO	56
3.3	PROCEDIMENTOS METODOLÓGICOS	58
3.4	MINERAÇÃO DE TEXTOS	61

3.4.1	SELEÇÃO DOS DADOS	61
3.4.2	PRÉ-PROCESSAMENTO	61
3.4.3	SELEÇÃO DE ATRIBUTOS.....	64
3.4.4	CLASSIFICAÇÃO	65
3.4.5	AVALIAÇÃO DO DESEMPENHO	67
3.4.6	RESULTADOS.....	68
3.5	CONSIDERAÇÕES FINAIS	72
3.6	REFERÊNCIAS	74
4	ARTIGO 3: Detecção de Infecções do Sítio Cirúrgico por meio da mineração de textos no Hospital de Clínicas de Porto Alegre.....	77
4.1	INTRODUÇÃO.....	77
4.2	REFERENCIAL TEÓRICO	79
4.3	PROCEDIMENTOS METODOLÓGICOS	81
4.4	MINERAÇÃO DE TEXTOS	84
4.4.1	SELEÇÃO DOS DADOS	84
4.4.2	PRÉ-PROCESSAMENTO	85
4.4.3	CLASSIFICAÇÃO	86
4.4.4	AVALIAÇÃO DO DESEMPENHO	86
4.4.5	RESULTADOS.....	88
4.5	CONSIDERAÇÕES FINAIS	91
4.6	REFERÊNCIAS	93
5.	CONSIDERAÇÕES FINAIS	98

1 INTRODUÇÃO

A qualidade do atendimento em saúde, de acordo com o comitê do Institute of Medicine (IOM), é definida como o grau em que os serviços de saúde para indivíduos e populações aumentam a probabilidade de resultados desejados e são consistentes com o conhecimento do profissional atual. Essa definição salienta a importância de resultados dos serviços de saúde, além de reconhecer as limitações do desempenho de um profissional de saúde e a existência de um componente estocástico nos processos assistenciais (IOM, 1990).

Os resultados de um sistema de saúde de qualidade podem ser percebidos a partir de seis dimensões: segurança do paciente, foco no paciente, eficácia, eficiência, oportunidade e equidade (IOM, 2001). Analisando a segurança do paciente, tema desta dissertação, é possível defini-la de maneira simples como sendo a prevenção de erros e efeitos adversos aos pacientes associados aos cuidados de saúde (WHO, 2017). A área de qualidade e, mais especificamente, a área de segurança do paciente fazem parte do tema desta dissertação a partir da análise de eventos adversos.

Eventos Adversos (EA) estão entre os resultados não desejados do atendimento em saúde e podem ser entendidos como danos provocados ao paciente pela assistência prestada que não são causados pela própria doença (HIATT *et al.*, 2000). Da mesma forma, Griffin e Resar (2009) afirmam que esses eventos podem gerar danos temporários, permanentes ou até mesmo levar à morte do paciente. Eventos inesperados e não desejados podem ocorrer em qualquer ambiente onde os cuidados de saúde são prestados (WHO, 2017).

Infecções em Sítio Cirúrgico (ISC) estão entre os EA mais comuns em pacientes hospitalizados, causando substancial aumento na mortalidade, taxas de reinternação e custos assistenciais (WATCHER, 2013; STONE *et al.*, 2009). O conhecimento exato das taxas de ISC a partir da sua detecção como EA pode ser usado como parte do mecanismo de feedback para diminuir a incidência futura de infecções cirúrgicas (MICHELSON, 2005).

A detecção de EA faz parte das tarefas de vigilância e melhoria da segurança do paciente nos hospitais. A partir da detecção, é possível realizar a análise de possíveis causas dos EA e tomar ações de prevenção e melhoria dos processos assistenciais (ZAPT e REASON, 1994). Nos últimos anos, vários bancos de dados

eletrônicos do ambiente hospitalar vêm sendo usados para realizar vigilância de EA, devido ao aumento das formas de coleta e armazenamento de informações de pacientes (FREEMAN *et al.*, 2013).

Entre os métodos que vêm sendo citados na literatura para apoiar a vigilância de EA estão os métodos de mineração de dados (MD), que aplicada a dados textuais não estruturados é chamada de Mineração de Textos (MT). Assim como a MD, a MT tem o princípio de extrair padrões não triviais, úteis e de interesse do usuário (TAN, 1999). O texto clínico pode ter informações essenciais para a vigilância; no entanto, métodos tradicionais de detecção podem utilizar recursos de maneira intensiva e demorada (FREEMAN *et al.*, 2013).

Esta dissertação teve como cenário a necessidade de identificação de métodos para apoiar ações de melhoria da segurança do paciente no Hospital de Clínicas de Porto Alegre (HCPA) a partir da previsão e detecção de infecções cirúrgicas no período pós-operatório. O HCPA faz parte de uma rede de hospitais universitários vinculados ao Ministério da Educação do Brasil, sendo que até o ano de 2015, 30 (trinta) hospitais universitários brasileiros já se tornaram usuários do Aplicativo de Gestão Hospitalar (AGH) (MEC, 2017), utilizado como fonte de dados para esta dissertação. A presente pesquisa foi desenvolvida no HCPA por este ser referência entre os hospitais universitários brasileiros e ser vinculado academicamente a Universidade Federal do Rio Grande do Sul (UFRGS), além de ter como missão o incentivo a pesquisas e à geração de conhecimento (HCPA, 2017a).

1.1 TEMA E OBJETIVOS

Esta dissertação tem como tema a segurança do paciente, abordando o uso de métodos de MT no contexto de um hospital universitário brasileiro. A discussão sobre métodos de MT também abrange estratégias de Processamento de Linguagem Natural (PLN) e Aprendizagem de Máquina (AM), de maneira que esses tópicos são tratados como parte do processo de MT.

O objetivo geral deste trabalho é avaliar o desempenho de métodos de mineração de textos na previsão e detecção de EA. Como objetivos específicos, este trabalho contempla:

- a) Identificar as áreas da saúde e métodos de mineração de textos utilizados na previsão e detecção de EA a partir de uma revisão sistemática da literatura;
- b) Avaliar o desempenho de métodos de mineração de textos na previsão de Infecções do Sítio Cirúrgico;
- c) Avaliar o desempenho de métodos de mineração de textos na detecção de Infecções do Sítio Cirúrgico.

Esses objetivos específicos correspondem aos três artigos apresentados nesta dissertação.

1.2 JUSTIFICATIVA DO TEMA E OBJETIVOS

Identificar e medir EA são tarefas fundamentais nos processos de saúde, formando uma base para a segurança do paciente (CLASSEN *et al.*, 2011). Segundo a Agência Nacional de Vigilância Sanitária ANVISA (2013), a ocorrência de EA pode acarretar no aumento na morbidade, na mortalidade, no tempo de tratamento dos pacientes e nos custos assistenciais, além de repercutir em outros campos da vida social e econômica da população.

Cerca de uma em 30 (trinta) cirurgias “limpas” será complicada por uma ISC. A taxa é significativamente maior em cirurgias “contaminadas”, de emergência ou em procedimentos prolongados e para pacientes com comorbidades clínicas (WATCHER, 2013). As ISC estão entre os EA mais comuns em pacientes hospitalizados, causando substancial aumento na mortalidade, taxas de reinternação e custos assistenciais (WATCHER, 2013).

No HCPA, a Comissão de Controle de Infecção Hospitalar (CCIH) executa ações buscando a melhoria da segurança do paciente, incluindo pesquisas entre as quais o presente estudo se encontra. São monitorados processos que apresentam potencial impacto sistêmico, realizada a vigilância das infecções, instituídas medidas de precaução, auditoria da prescrição e elaboração de política de uso antimicrobianos da instituição (HCPA, 2017b). Dessa forma, o controle de ISC no período pós-operatório é feito a partir de critérios que apontam a ocorrência de infecção, mas, apesar de esses critérios direcionarem o trabalho de vigilância, podem não levar a detecção de todas as infecções existentes.

Sistemas de detecção de eventos reduzem o custo da revisão de prontuários, identificando os casos que são mais apropriados para uma revisão detalhada (JHA *et al.*, 1998). Muitos dos dados clínicos presentes na forma narrativa são convenientes para expressar conceitos e eventos entre os envolvidos (pacientes e profissionais da saúde), porém dificultam a busca, a sumarização, o apoio à tomada de decisão e a análise estatística (MEYSTRE *et al.*, 2008).

Nesse cenário, a MT e o PLN, mais precisamente as técnicas de extração e recuperação da informação, têm apresentado potencial para apoiar a extração de conhecimento relacionado à ISC, antes detectados somente por meio de revisões manuais de registros hospitalares pós-operatórios. Da mesma maneira, a MT pode ser aplicada para apoiar a previsão de ISC a partir da descrição cirúrgica no período intra-operatório, conforme Figura 1.



Figura 1. Processo cirúrgico.

Nesse contexto, o resultado desta pesquisa, a partir dos objetivos propostos, servirá para que hospitais possam aplicar diferentes métodos de MT para apoiar o direcionamento da vigilância e controle de infecções. Além disso, os resultados desta pesquisa podem servir para que outras aplicações em diferentes tipos de EA venham a ser realizadas, proporcionando a melhoria da qualidade em diferentes áreas da saúde.

1.3 DELINEAMENTO DO ESTUDO

Uma vez definidos os objetivos e apresentada a relevância do tema desta dissertação, se faz necessário descrever o delineamento pelo qual os objetivos foram alcançados. Esta seção descreve os métodos de pesquisa e de trabalho utilizados neste estudo.

1.3.1 MÉTODO DE PESQUISA

Esta dissertação é caracterizada como sendo de natureza aplicada, pois apresenta a revisão e aplicação de métodos de MT buscando soluções para problemas específicos da área da saúde. Do ponto de vista da abordagem, esta pesquisa configura-se como qualitativa pela seleção e análise das publicações, mas também quantitativa quando lida com a avaliação de desempenho de métodos de MT (GIL, 2008; CERVO e BERVIAN, 2002).

Da mesma forma, o objetivo geral aqui proposto possui caráter descritivo. Essa caracterização ocorre pelo intuito de descrever a forma com que métodos de MT já apresentados na literatura podem ser aplicados na previsão e detecção de EA (GIL, 2008).

1.3.2 MÉTODO DE TRABALHO

O desenvolvimento deste trabalho contemplou a realização de três etapas. A primeira etapa diz respeito à revisão sistemática da literatura, que buscou identificar as áreas da saúde e os métodos de MT utilizados para a detecção de EA. As publicações foram selecionadas observando o critério de utilizar métodos de MT para detectar ou prevenir EA.

A segunda etapa consiste na avaliação de métodos de MT quanto ao seu potencial para prever ISC. Inicialmente, por meio da literatura, são apresentados aspectos específicos da prevenção de infecções. A partir disso é apresentada uma aplicação de MT buscando prever ISC utilizando o texto de descrições cirúrgicas realizadas no HCPA. O processo de MT aqui utilizado contempla as atividades de seleção dos dados, pré-processamento de texto, aplicação de métodos de classificação e por fim a avaliação do desempenho.

Após a previsão de ISC, se faz necessária a análise de métodos para encontrar infecções no local cirúrgico a partir de evoluções dos pacientes no período pós-operatório. Nesse contexto, a terceira etapa diz respeito à avaliação de métodos de mineração de textos quanto ao seu potencial para detecção de ISC. São utilizados dados textuais de evoluções dos pacientes referentes aos 30 (trinta) dias após uma cirurgia, como anotações da internação ou reconsulta. Assim como na segunda etapa,

o processo de mineração de textos contempla as atividades de seleção dos dados, pré-processamento de texto, aplicação de métodos de classificação e avaliação de desempenho.

1.4 DELIMITAÇÕES DO ESTUDO

Este trabalho contempla a análise da literatura sobre EA e MT em geral e a aplicação de MT na previsão e detecção de EA em um hospital universitário brasileiro. Os dados textuais analisados na segunda e terceira etapas estão no idioma português e dizem respeito às descrições cirúrgicas e evoluções de pacientes até 30 (trinta) dias após cirurgias.

Não foram utilizados filtros de pacientes por condição clínica ou patologia. Este estudo contemplou apenas a análise do texto livre; sendo assim, não foram utilizados atributos como resultados de exames ou outras informações do paciente que não estão presentes no texto. Também não foram realizadas análises semânticas ou gramaticais de texto, de maneira que apenas a frequência de conjuntos de até três palavras foi utilizada.

As aplicações de mineração de textos aqui descritas foram realizadas na linguagem de programação Python versão 3.5 (PYTHON, 2017), escolhida por ser gratuita e viabilizar futuras implementações práticas deste estudo por outros pesquisadores. O Python também possui bibliotecas de AM e processamento de dados textuais, Natural Language Toolkit (NLTK), Scikit-Learn, nas quais o presente estudo foi desenvolvido (BIRD *et al.* 2009; PEDREGOSA *et al.*, 2011).

1.5 ESTRUTURA DA DISSERTAÇÃO

Esta dissertação está organizada em cinco capítulos. O primeiro capítulo aborda a introdução do trabalho e os objetivos, apresentando os métodos de pesquisa e o delineamento do estudo. Nos capítulos 2, 3 e 4, são apresentadas as etapas nas quais desta dissertação foi desenvolvida no formato de artigos. O primeiro artigo apresenta uma revisão sistemática da literatura, a partir da qual é possível perceber as áreas da saúde e os métodos que utilizam, bem como as áreas e métodos que

relatam maior necessidade de estudos. A seguir, no segundo e terceiro artigos, aplicações de MT são apresentadas, expondo os melhores métodos para previsão e detecção de ISC no período pós-operatório no HCPA. Por fim, são apresentadas as considerações finais no capítulo 5.

1.6 REFERÊNCIAS

ANVISA. Critérios Diagnósticos de Infecções Relacionadas à Assistência à Saúde. Segurança do Paciente e Qualidade em Serviços de Saúde. Brasília, 2013.

CERVO, A. L. E., BERVIAN, P. A. Metodologia Científica. 5ª edição. Prentice Hall, São Paulo. 2002.

CLASSEN, D. C., RESAR, R., GRIFFIN, F., FEDERICO, F., FRANKEL, T., KIMMEL, N., WHITTINGTON, J. C., FRANKEL, A., SEGER, A., JAMES, B. C. 'Global trigger tool' shows that adverse events in hospitals may be ten times greater than previously measured. Health affairs, 2011.

FREEMAN, R., MOORE, L. S., ÁLVAREZ, L. G., CHARLETT, A., & HOLMES, A. Advances in electronic surveillance for healthcare-associated infections in the 21st Century: a systematic review. Journal of Hospital Infection, 2013.

GIL, A. C. Métodos e Técnicas de Pesquisa Social. 6ª edição, Atlas, São Paulo. 2008.

HOSPITAL DE CLÍNICAS DE PORTO ALEGRE - HCPA. Disponível em: <https://www.hcpa.edu.br/content/view/13/2616/>. Acessado em 20/02/2017, a.

HOSPITAL DE CLÍNICAS DE PORTO ALEGRE - HCPA. Disponível em: <https://www.hcpa.edu.br/content/view/7925/2429/>. Acessado em: Acessado em: 18/02/2017, b.

HIATT, H. H., BARNES, B. A., BRENNAN, T. A., LAIRD, N. M., LAWTHERS, A. G., LEAPE, L. L., LOCALIO, A. R., NEWHOUSE, J. P., PETERSON, L. M., THORPE, K. E. A study of medical injury and medical malpractice. An overview. New England Journal of Medicine, 1989.

IOM - INSTITUTE OF MEDICINE. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, D.C: National Academy Press; 2001.

IOM - INSTITUTE OF MEDICINE. Medicare: A Strategy for Quality Assurance. Vol.1. Washington, DC: National Academy Press, 1990

JHA, A.K., KUPERMAN, G.J., TEICH, J.M., LEAPE, L., SHEA, B., RITTENBERG, E., et al. Identifying adverse drug events: development of a computer- based monitor and comparison with chart review and simulated voluntary report. Journal American Medical Informatics Association, 1998.

MEYSTRE, S.M. SAVOVA, G.K. KIPPER-SCHULER, K.C. HURDLE, J.F. Extracting information from textual documents in the electronic health record: a review of recent research. Yearbook Medical Informatics. 2008.

MEC - MINISTÉRIO DA EDUCAÇÃO. Disponível em: <http://portal.mec.gov.br/hospitais-universitarios/aghu>. Acessado em: 23/04/2017.

MICHELSON, J. D.; PARISEAU, J. S.; PAGANELLI, W. C. Assessing surgical site infection risk factors using electronic medical records and text mining. *AJIC: American Journal of Infection Control*, 2014.

PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., MICHEL, V., THIRION, B., GRISEL, O., BLONDEL, M., PRETTENHOFER, P., WEISS, R., DUBOURG, V., VANDERPLAS, J., PASSOS, A., COURNAPEAU, D., BRUCHER, M., PERROT, M., DUCHESNAY, E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 2011.

PYTHON SOFTWARE FOUNDATION. Disponível em: <https://www.python.org/>. Acessado em: 29/01/2017.

STONE, P. W., KUNCHES, L., HIRSCHHORN, L. Cost of hospital-associated infections in Massachusetts. *American journal of infection control*, 2009.

TAN, A. Text Mining: The state of the art and the challenges. *Proceedings of the PAKDD 1999 Workshop on Knowledge Discovery from Advanced Databases*, 1999.

WATCHER, R. M. *Compreendendo a Segurança do Paciente – 2ª ed.* Mcgraw-Hill Education. 2013.

WHO - WORLD HEALTH ORGANIZATION. Disponível em: <http://www.euro.who.int/en/health-topics/Health-systems/patient-safety/patient-safety>. Acessado em: 29/01/2017.

ZAPT, D., REASON, J.T. Introduction to error handling. *Applied Psychology*, 1994.

2 ARTIGO 1: Detecção de eventos adversos por meio da mineração de textos: uma revisão sistemática da literatura

Resumo: O desenvolvimento tecnológico verificado na área da saúde tem motivado novas pesquisas para a detecção de Eventos Adversos (EA). Identificar e medir a ocorrência de eventos que causam danos aos pacientes é fundamental para a melhoria da segurança. Em hospitais que utilizam sistemas informatizados e prontuários eletrônicos, o uso de técnicas de mineração de textos pode ser uma alternativa viável para a melhoria de aspectos da segurança da assistência médica. Este estudo teve por objetivo identificar as áreas da saúde e métodos de Mineração de Textos (MT) utilizados na detecção e prevenção de EA a partir de uma revisão sistemática da literatura. A busca foi realizada a partir do Portal de Periódicos da Capes, o qual permitiu o acesso as bases de literatura ACM Digital Library, CiteSeerX, IEEE Xplore, Pubmed e ScienteDirect. Foram selecionadas publicações que relatam a utilização de métodos de MT aplicados na prevenção e detecção de EA. Ao todo, 54 publicações fizeram parte deste estudo, nas quais foi constatada a predominância da farmacovigilância no uso de métodos de MT. Quanto aos métodos utilizados, foi percebido que as pesquisas buscam propor novos métodos e sistemas de mineração para áreas específicas da saúde; no entanto, métodos tradicionais de MT como Suport Vector Machines, classificadores Naive Bayes e algoritmos de associação também são apresentados.

Palavras-chave: eventos adversos; mineração de textos; revisão sistemática da literatura.

2.1 INTRODUÇÃO

Eventos Adversos (EA) podem ser entendidos como os danos provocados ao paciente pela assistência prestada, que não são causados pela própria doença (HIATT *et al.*, 2000). Essa definição inclui todos os aspectos do tratamento médico, incluindo os sistemas e equipamentos utilizados para fornecer a assistência médica e também a incapacidade de fornecer diagnóstico e tratamento ao paciente (WHO,

2005). Alguns danos são inerentes ao tratamento; porém, os danos preveníveis, segundo a definição comum encontrada na literatura, são aqueles que possuem uma causa identificável e modificável (NABHAN *et al.*, 2012).

Os EA podem ocorrer em função de diversas causas, dentre elas o uso de medicamentos (WHO, 2005). Boa parte da literatura recente tem buscado identificar Reações Adversas a Medicamentos (RAM) por meio da extração de informações de textos clínicos, aplicando técnicas estatísticas em relatórios espontâneos e ensaios clínicos (HARPAZ *et al.*, 2014).

Prontuários eletrônicos contêm as informações do estado do paciente, prescrições de medicamentos e os procedimentos nos quais os pacientes são submetidos. Dessa forma, os prontuários possuem evidências ou pistas da ocorrência de EA, as quais podem causar danos temporários que requerem alguma intervenção inicial, prolongada, ou até mesmo permanente, bem como danos que não podem ser completamente tratados ou levam à morte do paciente (GRIFFIN e RESAR, 2009).

A identificação e posterior análise de EA é importante para identificar aspectos dos serviços de saúde propensos a falhas. A crescente adoção de softwares que apoiam os serviços de saúde torna necessária a utilização de técnicas e ferramentas que apoiem a extração de conhecimento útil das bases de dados dos hospitais. Novas tecnologias permitem o armazenamento de uma quantidade cada vez maior de dados dos processos hospitalares; no entanto, dados por si só carecem de conhecimento do negócio para que possam agregar valor à organização (FAYYAD, 1996).

A Mineração de Dados (MD) é uma área que tem tido muitos avanços nos últimos anos em função das evoluções de hardware e software, que vem viabilizando a análise de diferentes tipos de dados, como os dados textuais (AGGARWAL e ZHAI, 2012). A Mineração de Texto (MT), que é uma variação da MD, é uma importante disciplina do Processamento de Linguagem Natural (PLN) (BIEMANN e MEHLER, 2015).

Muitos dos dados clínicos presentes na forma narrativa são convenientes para expressar conceitos e eventos entre os envolvidos (pacientes e profissionais da saúde), porém dificultam a busca, a sumarização, o apoio à tomada de decisão e a análise estatística (MEYSTRE *et al.*, 2008). Nas áreas da saúde e biomedicina, conhecimentos são disseminados em forma textual a partir de literatura científica,

técnica, relatórios administrativos, da mesma forma que as informações de pacientes estão contidas em notas narrativas e relatórios (SHORTLIFFE e CIMINO, 2014).

Nesse contexto, o PLN, mais precisamente as técnicas de extração e recuperação da informação, tem apresentado potencial para apoiar a extração de conhecimento relacionado a EA. Para conduzir pesquisas à melhoria da segurança na saúde é necessário identificar os métodos atuais para detecção de eventos que causam danos aos pacientes e as áreas da saúde nas quais esses métodos vêm sendo aplicados. Nesse sentido, este artigo tem por objetivo identificar métodos de MT e as áreas da saúde em que esses métodos estão sendo aplicados na detecção e prevenção de EA, a partir de uma revisão sistemática da literatura.

O presente artigo está organizado em três seções além dessa introdução. Na seção 2.2 são apresentados os procedimentos metodológicos utilizados e a seção 2.3 apresenta a revisão sistemática da literatura. Na seção 2.4 são apresentadas as considerações finais deste estudo.

2.2 PROCEDIMENTOS METODOLÓGICOS

Este trabalho apresenta uma revisão das pesquisas sobre mineração de textos de registros eletrônicos de saúde para detecção de EA. Este trabalho possui natureza básica e utiliza abordagens qualitativas na seleção das publicações, mas também quantitativas no que diz respeito à análise das áreas e métodos de MT utilizados. Após a identificação da necessidade de uma revisão sistemática do assunto, foi definido o local de busca, as palavras-chave, os critérios de inclusão e o protocolo de leitura.

A bibliografia desta pesquisa foi selecionada por meio do Portal de Periódicos da Capes (PPC) a partir da busca por publicações até junho de 2016, utilizando os seguintes termos em inglês: “text mining” ou “natural language processing” e “adverse event(s)”, contidos no título, resumo ou no corpo das publicações. Também foram selecionados trabalhos a partir do referencial bibliográfico das publicações selecionadas no PPC.

O PPC centraliza o acesso a diversas bases de dados internacionais. Não foram utilizados filtros por bases de dados específicas, sendo que os artigos selecionados foram encontrados nas seguintes bases: ACM Digital Library, CiteSeerX,

IEEE Xplore, Pubmed e ScienteDirect. Ao todo 54 publicações foram selecionadas, conforme Figura 2 (MOHER et al., 2009).

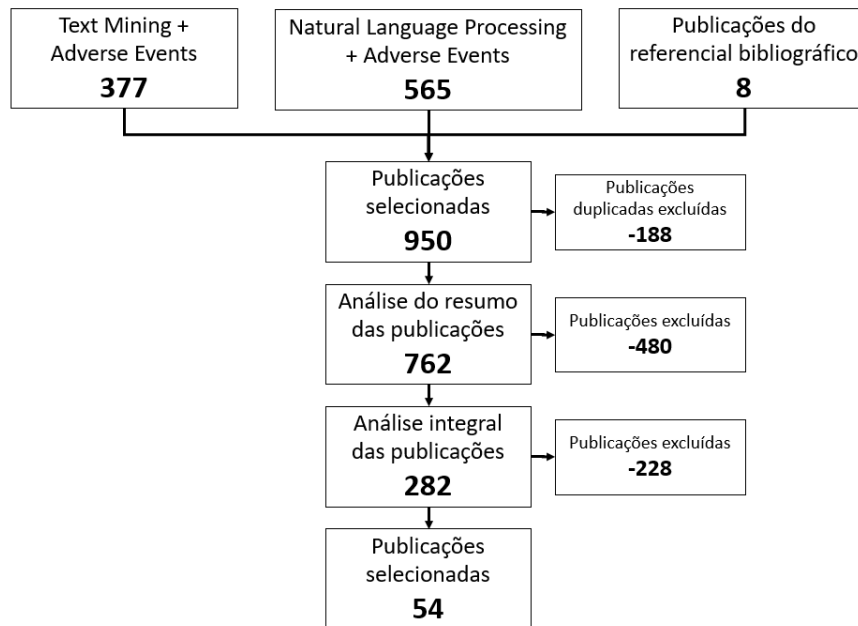


Figura 2. Publicações selecionadas.

A partir da análise do título, resumo e leitura integral do artigo, foram aplicados critérios de inclusão e exclusão quanto ao método e a área do estudo. Quanto ao método foram selecionadas publicações que utilizaram técnicas de mineração de textos. Já a área teve de estar relacionada, de maneira geral, a EA, incluindo também as reações adversas a medicamentos e vacinas. Na etapa seguinte, foi realizada a leitura integral de cada estudo buscando atender as questões de pesquisa e avaliando a qualidade das publicações, conforme questionário do Apêndice A (RADJENOVIC *et al.*, 2013), verificando se as publicações atendem totalmente, parcialmente ou não atendem as questões de qualidade.

2.2.1 QUESTÕES DE PESQUISA

A condução de uma revisão sistemática requer a elaboração de questões de pesquisa que direcionem os caminhos metodológicos a serem seguidos (INSTITUTE OF MEDICINE, 2011). As questões de pesquisa foram elaboradas em concordância com estudos anteriores que apresentam revisões da literatura envolvendo MT e EA

(GOVINDAN et al., 2010; MEYSTRE et al., 2008; DEMNER-FUSHMAN et al., 2009; HARPAZ, et al., 2014).

O estabelecimento de questões de pesquisa é importante não apenas para os pesquisadores, mas também para profissionais que buscam soluções para a melhoria da segurança do paciente. Para esta revisão sistemática da literatura, as questões de pesquisa apresentadas na Tabela 1 foram formuladas:

Tabela 1. Questões de Pesquisa

1	Em que áreas da saúde a MT tem sido utilizada?	Identificar quais áreas da saúde mais têm buscado soluções envolvendo mineração de textos para prevenir e detectar EA.
2	Quais técnicas, algoritmos ou sistemas são usados na MT?	Identificar quais técnicas, algoritmos ou sistemas são mais utilizados para prevenir e detectar EA em saúde.

2.3 REVISÃO SISTEMÁTICA DA LITERATURA

2.3.1 TRABALHOS CORRELATOS

Foram considerados trabalhos correlatos aqueles que se dedicaram a revisão da literatura sobre EA e MT. Dessa forma, foram selecionadas quatro pesquisas que apresentam diferentes abordagens e escopo de revisão sobre o assunto abordado.

Govindan *et al.* (2010) desenvolveram uma revisão sistemática da literatura buscando identificar, descrever e avaliar a eficácia dos métodos automatizados de detecção de danos aos pacientes. Foram encontradas 4 pesquisas relacionadas ao processamento de linguagem natural e outras 39 relacionadas a outros métodos. Entre as pesquisas, 21 envolviam a detecção de RAM, 10 buscaram detectar EA em geral, 8 buscaram infecções hospitalares e 4 buscaram outros EA específicos.

Já Meystre *et al.* (2008) buscaram na literatura recente (PubMed, ACM Digital Library e anais de congressos) trabalhos que envolvessem a extração de informações de Registros Eletrônicos Hospitalares (REH). Os autores apresentaram uma visão geral dos avanços quanto ao pré-processamento de documentos textuais, análise e detecção de atributos contextuais, extração de códigos e mineração de textos clínicos,

enriquecimento de REH para apoio à tomada de decisões, extração de informações para vigilância médica, apoio à pesquisa, gerenciamento automático de terminologias, corpora de textos clínicos e sua anotação e mineração de textos clínicos.

Demner-fushman *et al.* (2009) tiveram como foco em sua revisão da literatura analisar os avanços de métodos de PLN em sistemas clínicos de apoio à decisão. A análise ocorreu a partir de modelos, técnicas como Reconhecimento de Entidades Nomeadas (REN), extração de contexto e uma visão dos sistemas atuais de apoio à decisão em cenários especializados, como relatórios de radiologia, emergência e patologia. Os autores afirmam que é preciso desenvolver métodos de avaliação para medir o impacto das soluções de PLN.

Harpaz *et al.* (2014) fazem uma análise da literatura buscando identificar os avanços das técnicas da mineração de textos na farmacovigilância. A revisão é organizada a partir das fontes de dados textuais que podem contribuir para a farmacovigilância, como literatura biomédica, narrativas clínicas, rótulos de produtos, mídias sociais e registros de eventos de buscas na internet. Como conclusões, os autores citam que o acesso limitado a determinadas fontes de dados, que também não foram projetadas para a farmacovigilância, formam as principais dificuldades para a extração de informações de textos na área.

O presente estudo difere da revisão de Govindan *et al.* (2010) no sentido de que a revisão dos autores não é direcionada apenas às tecnologias usadas para a extração de informações de texto livre. Já Meystre *et al.* (2008) e Demner-fushman *et al.* (2009) diferem da presente pesquisa quanto às fontes onde foram realizadas as buscas e ao levantamento das áreas da saúde onde as técnicas são aplicadas. Por fim, Harpaz *et al.* (2014) diferem deste estudo por terem como foco apenas a farmacovigilância. De maneira geral, o presente estudo difere das pesquisas anteriores em relação ao escopo e objetivo da revisão.

2.3.2 MINERAÇÃO DE TEXTOS

A MT tipicamente compreende dois ou três dos seguintes passos: recuperação da informação, extração da informação e, por fim, a mineração de dados (HEARST, 1999). Devido à complexidade dos eventos médicos, a extração efetiva de eventos normalmente requer uma análise a partir da estrutura das sentenças

(FRIEDMAN e ELHADAD, 2014), o que pode ocasionar na necessidade de utilização de métodos complementares de análise.

Nesse sentido, das 54 publicações analisadas, 32 utilizaram mais de um método de mineração de textos para detectar ou apoiar a identificação de EA, totalizando 125 métodos, técnicas ou sistemas utilizados. Ao todo, 79% desses métodos dizem respeito a novas soluções específicas propostas pelos autores, sendo que 16% correspondem aos métodos Suport Vector Machine (SVM), classificadores Naive Bayes e Association Rule Mining e 5% aos métodos Decision Tree, Nearest Neighbor, Logistic Regression e Clustering.

Kang *et al.* (2012) desenvolveram um método *ensemble* para melhorar a extração de conceitos de dados clínicos. Ensemble systems, também chamados de sistemas multiclassificadores, combinam a saída de diferentes métodos de classificação e apresentam o resultado das melhores combinações, esperando obter um resultado melhor do que o melhor classificador individual (POLIKAR, 2006).

Juntamente com a mineração de textos, o PLN tem sido usado com sucesso em vários domínios específicos da medicina (BAUD *et al.*, 1995; SAGER *et al.*, 1995; VILAR *et al.* 2012), tais como: (i) infecções nosocomiais (BATES *et al.*, 2003; MENDONÇA *et al.*, 2005); (ii) reações adversas a medicamentos (ARAMAKI *et al.*, 2010; BIAN *et al.*, 2012; CHEE *et al.*, 2011; HAERIAN *et al.*, 2012; WANG *et al.*, 2014); (iii) RAM após vacinas (BAER *et al.*, 2016; HUR *et al.*, 2012); (iv) e EA de maneira geral (HAZLEHURST *et al.* 2005; LI *et al.* 2014; ROCHEFORT *et al.* 2015a; ROCHEFORT *et al.* 2015b).

O PLN refere-se aos métodos automatizados que convertem dados de texto livre em formato computacional compreensível (ALLEN, 1995). Não está claro, porém, se processadores de linguagem natural podem detectar uma ampla gama de EA complexos com precisão suficiente para ajudar as instituições de saúde de maneira significativa (MELTON e HRIPCSAK, 2005).

Técnicas de PNL têm sido divididas em duas grandes categorias: simbólicas e estatísticas. Técnicas simbólicas, ou gramaticais, usam as características da linguagem, isto é, a semântica, a sintaxe, e as relações entre frases, para interpretar um documento narrativo com objetivo de classificá-lo em um conjunto de categorias discretas (CHAPMAN, 2006). Já técnicas estatísticas de PLN usam a distribuição de palavras e a frequência de frases para classificar automaticamente um conjunto de documentos em um conjunto discreto de categorias predefinidas (MITCHELL, 1997).

Foram encontrados estudos que utilizam PLN para apoiar seus métodos de detecção de EA e outros com o objetivo de avaliar a eficácia da sua utilização no contexto médico. Por exemplo, Melton e Hripcsak (2005) tiveram como objetivo determinar se o PLN pode detectar efetivamente EA definidos no Sistema de Relatórios e Rastreamento de Ocorrências em Pacientes de Nova York (NYPORTS) usando resumos de alta. De 57.452 resumos de alta eletrônicos, o sistema identificou 1.590 eventos em 1.461 casos e a revisão manual identificou 704 eventos em 652 casos.

Ong *et al.* (2012) buscaram explorar a viabilidade da utilização de classificação estatística de textos para detectar automaticamente os eventos de alto risco em relatórios de incidentes clínicos. De maneira geral, os autores tiveram maior precisão da previsão utilizando um conjunto de dados especializado, em comparação com o conjunto de dados com várias especialidades médicas.

Além do PLN, as principais abordagens encontradas envolvem a classificação das sentenças dos textos médicos, utilizando técnicas de aprendizagem de máquina como os algoritmos de classificação Support Vector Machine (SVM) e Naive Bayes (NB), que são comumente usados na classificação de documentos (SEBASTIANI, 1998; BRUIJN *et al.*, 2011). Na mesma direção, Genkin *et al.* (2007) apresentam a aplicação de uma abordagem de classificação a partir de regressão logística bayesiana em um conjunto de dados do MedLine, que pode ser usada para detecção de RAM.

2.3.3 DETECÇÃO DE EVENTOS ADVERSOS

Gestores e profissionais de saúde precisam de tecnologia da informação associada a melhores formas de coleta de dados para melhorar a segurança do paciente (INSTITUTE OF MEDICINE, 2001). Sistemas informatizados para a detecção de eventos contam com alertas, tanto no caso de eventos iminentes, para prevenção, quanto em eventos ocorridos, para gerenciamento (BATES *et al.*, 2003). Sistemas de detecção de eventos reduzem o custo da revisão de prontuários, identificando os casos que são mais apropriados para uma revisão detalhada (JHA *et al.*, 1998).

A literatura tem abordado a detecção de EA de duas áreas principais. A primeira diz respeito aos EA de maneira geral, descritos como danos ou ocorrências

desfavoráveis aos pacientes provenientes de diversos fatores que podem variar de acordo com a patologia analisada. Já a segunda área está relacionada especificamente ao uso de medicamentos.

Evento adverso é qualquer ocorrência médica desfavorável, em um paciente ou assunto de investigação clínica, administrando um produto farmacêutico, o qual não possui necessariamente uma relação causal com este tratamento.

Um evento adverso pode, portanto, ser qualquer sinal desfavorável e não intencional (incluindo uma evidência laboratorial anormal, por exemplo), sintoma ou doença temporariamente associada ao uso de um medicamento, quer sejam ou não considerados relacionados ao medicamento. (International Council for Harmonisation, 1994, pág. 2, tradução do autor).

Por definição, a análise de EA abrange também o estudo dos eventos adversos a medicamentos, também descritos como reações ou efeitos adversos a medicamentos, dado que a possibilidade de uma relação causal entre o medicamento e o efeito adverso não pode ser descartada.

Reação adversa a medicamento refere-se a toda a resposta nociva e não desejada relacionada a um medicamento em qualquer dose. A frase "respostas a medicamentos" significa que uma relação causal entre um medicamento e um evento adverso é, pelo menos, uma possibilidade razoável de, ou seja, a relação não pode ser descartada. (International Council for Harmonisation, 1994, pág. 2, tradução do autor).

Zapt e Reason (1994) afirmam que a detecção de EA representa uma oportunidade de aprendizado por meio de uma perspectiva cognitiva, indicando fatores que cercam os eventos e podem ser melhorados. Vários estudos mostram que relatórios de incidentes apoiam a identificação de problemas evitáveis, além de fornecerem informações úteis sobre o contexto em que os EA ocorrem, em comparação com os métodos tradicionais, tais como revisão de prontuários médicos e registros eletrônicos hospitalares (BECKMANN *et al.*, 2003; HONIGMAN *et al.*, 1993; LEPENDU *et al.*, 2013; O'NEIL *et al.*, 1993).

Nas publicações analisadas, 34 (63%) utilizaram bases de dados públicas, o que possibilita a verificação do método proposto e também a repetibilidade do estudo por outros pesquisadores. As fontes das informações, sejam elas prontuários ou relatórios de incidentes, podem ter diferente desempenho para cada tipo de EA. Tinoco *et al.* (2011) tiveram como objetivo entender como a fonte de informação afeta diferentes métodos de vigilância de EA. Ao comparar um método computadorizado e um método manual utilizando narrativas clínicas, foram encontradas mais infecções

hospitalares por meio do método computadorizado, 92%, contra 34% do método manual, porém, ambos os métodos tiveram resultado similar na detecção de RAM, 52% e 51%.

EA podem ser abordados de maneira bastante ampla, tendo diferentes características para cada patologia. Dessa forma, as abordagens de mineração de textos encontradas na literatura são aplicadas a diversos tipos de EA. Dentre as especialidades médicas abordadas na literatura está a área cirúrgica, direcionada principalmente à detecção de Infecções de Sítio Cirúrgico (ISC).

Campillo-Gimenez *et al.* (2013) utilizaram uma ferramenta chamada Nomindex na detecção de ISC nos 30 (trinta) dias seguintes a neurocirurgias. Também utilizando dados de um prontuário eletrônico integrado, o método de mineração de textos de Michelson *et al.* (2014) foi capaz de encontrar 22 ISC detectadas pela vigilância tradicional, bem como um adicional de 37 ISC que não foram detectadas pela vigilância tradicional. Registros de anestesia também podem ser fonte de informações para detecção de ISC. Eles forneceram informações sobre a administração de antibióticos, duração da cirurgia e uma pontuação da American Society of Anesthesiologists (ASA) para o risco da anestesia pré-operatória (MICHELSON *et al.*, 2014).

Outros eventos assistenciais não programados também podem servir como pista para a detecção de EA, como a readmissão hospitalar, avaliações no departamento de emergência, ou a própria internação não planejada, que são potenciais indicadores de práticas inseguras ou ineficientes (WEINICK *et al.*, 2010; SPINKS *et al.*, 2014). Para identificar eventos não programados, Tamang *et al.* (2015) analisaram 308.096 documentos com texto livre sobre câncer de mama, câncer gastrointestinal e câncer torácico utilizando métodos de mineração de texto, constatando que as técnicas de mineração e a combinação de relatórios aumentou a identificação de pacientes com um ou mais eventos não programados.

Lesões relacionadas com quedas também são um importante problema de saúde, especialmente em populações envelhecidas. O histórico de queda é um dos indicadores clínicos mais importantes para identificação de pacientes com alto risco de queda (GANZ *et al.*, 2007). McCart *et al.* (2012) e Rochefort *et al.* (2015a) tiveram resultados positivos ao verificarem quão bem métodos estatísticos de mineração de textos e PLN podem apoiar na prevenção de quedas de pacientes.

Métodos manuais para detecção de EA geralmente utilizam ferramentas de gatilho com regras que caracterizam um EA ou Erro Médico (EM). Um EM é definido como o fracasso de uma ação planejada para ser concluída conforme previsto ou o uso de um plano errado para alcançar um objetivo (SHAREK e CLASSEN, 2006; BATES *et al.*, 1995). A distinção entre EA e EM é importante porque muitos erros médicos não levam a danos, enquanto EA, por definição, estão ligados a danos ao paciente, mas podem não ser devidos a um EM (NEBEKER *et al.*, 2004). Nesse contexto, Li *et al.* (2014) desenvolveram uma abordagem para detecção de EA e EM por meio de fenotipagem em uma unidade de tratamento intensivo neonatal, concluindo que as técnicas de PLN utilizadas apresentaram maior sensibilidade do que os métodos de vigilância tradicionais.

Ross *et al.* (2013) também tiveram como objetivo avaliar técnicas de classificação de texto para melhorar a recuperação de informações em um banco de dados de fenótipos e genótipos (dbGaP). Foram utilizados algoritmos de aprendizagem de máquina treinados em textos de estudos sobre fenótipos e genótipos (PubMed) para identificação de metadados sobre coração, pulmão, e exames de sangue, podendo ser aplicados na detecção de EA.

Já o tromboembolismo venoso, que inclui a trombose venosa profunda e o embolismo pulmonar, é uma das complicações mais comuns de hospitalização e também pode ser considerado um EA (GEERTS *et al.*, 2008; SPYROPOULOS *et al.*, 2009). Rochefort *et al.* (2015b), usando técnicas estatísticas e PLN, buscaram determinar a acurácia de técnicas de PLN na detecção do tromboembolismo venoso, concluindo que o PLN pode apoiar nas tarefas de detecção em relatórios de narrativas de radiologia. Ainda, Rochefort *et al.* (2015a) tiveram um objetivo semelhante avaliando a acurácia de técnicas de PLN, porém aplicado na detecção de pneumonia adquirida no hospital, infecções da corrente sanguínea associadas ao cateter e quedas hospitalares, também obtendo resultados positivos.

Alguns estudos têm abordado aspectos específicos relevantes na detecção e análise de EA, como os aspectos temporais presentes em narrativas clínicas. Clark *et al.* (2014) tiveram o objetivo de ilustrar como a web semântica, a partir do Clinical Narrative Temporal Relation Ontology (CNTRO), pode ser usada para analisar as propriedades temporais de eventos documentados por meio de múltiplas narrativas médicas. Em narrativas clínicas, muitas características temporais são expressas de maneira relativa, por exemplo, “no início da semana passada” ou “na próxima sexta-

feira". Ignorar essas informações seria renunciar a uma valiosa informação que poderia apoiar a pesquisa de várias formas. Em relatos clínicos, a informação temporal muitas vezes não é expressa explicitamente, devendo ser inferida antes da análise dos dados (CLARK *et al.*, 2014).

Por fim, Hazlehurst *et al.* (2005) desenvolveram o sistema MediClass para apoiar a detecção de eventos em registros eletrônicos hospitalares. Os autores afirmam que o sistema pode apoiar operações clínicas e serviços de pesquisa por meio da avaliação de qualidade assistencial, vigilância de doenças e detecção de EA.

2.3.4 FARMACOVIGILÂNCIA

A detecção de EA decorrentes do uso de medicamentos é uma tarefa da farmacovigilância. A Organização Mundial da Saúde (OMS) define a farmacovigilância como "a ciência e as atividades relacionadas à detecção, avaliação, compreensão e prevenção de efeitos adversos ou outros possíveis problemas relacionados ao medicamento" (WHO, 2002).

A farmacovigilância é necessária para a identificação sistemática de associações causais entre drogas e efeitos colaterais, além de tomar ações corretivas, tanto para novos medicamentos lançados no mercado, quanto para os medicamentos já em uso. A farmacovigilância baseia-se na coleta de notificações espontâneas de relatórios de EA (YELESWARAPU *et al.*, 2014)

2.3.4.1 FARMACOVIGILÂNCIA EM REDES SOCIAIS

Uma abordagem relativamente nova para a farmacovigilância é o uso de texto não estruturado de sites relacionados à saúde para extrair associações de EA. Essa é uma consequência da recente tendência de pessoas em relatarem suas experiências pessoais em blogs e fóruns com mais frequência do que as relatam aos médicos. Sites relacionados à saúde permitem que as pessoas discutam suas condições médicas umas com as outras (YELESWARAPU *et al.*, 2014; YANG *et al.*, 2015; BENTON *et al.* 2011; SAMPATHKUMAR *et al.*, 2014; CHEE *et al.* 2011).

Nos últimos anos, fóruns na Web tem oferecido uma oportunidade para pacientes e médicos compartilharem informações de saúde sobre medicamentos. Hadzi-Puric e Grmusa (2012) tiveram como objetivo descobrir e quantificar associações entre medicamentos e reações adversas por meio de métodos estatísticos baseados em diferentes medidas de desproporcionalidade. Foram coletados dados de oito sites da internet, a fim de explorar a descoberta automática de reações adversas a medicamentos.

Comentários dos usuários de sites de redes sociais relacionados com a saúde também podem refletir RAMs conhecidas e potencialmente servir como um sistema de alerta sobre RAMs desconhecidas (HU e LIU, 2004). Nesse sentido, sistemas que extraem menções a RAM a partir de comentários de usuários em redes sociais vêm obtendo resultados positivos em tarefas de farmacovigilância (NIK FARJAM e GONZALEZ, 2011; LEAMAN *et al.*, 2010)

Sistemas de relatórios de EA tradicionais têm sido lentamente adaptados para relatórios de eventos online. Da mesma forma, o número de pacientes que se voltaram para as mídias sociais para compartilhar suas experiências com medicamentos, dispositivos médicos e vacinas também vem crescendo e com isso surge uma nova fonte de dados para a farmacovigilância e também para acompanhamento de medicamentos por órgãos reguladores. Nesse sentido, Freifeld *et al.* (2014) buscaram avaliar a concordância entre reações e notificações espontâneas mencionadas no Twitter e as recebidas por uma agência reguladora. Também tendo como fonte de dados comentários de usuários no Twitter, Bian *et al.* (2012) utilizaram técnicas de PLN e classificação de textos para encontrar potenciais EA em usuários de medicamentos.

2.3.4.2 REAÇÕES ADVERSAS A MEDICAMENTOS

Conforme mencionado anteriormente, RAM refere-se a toda a resposta nociva e não desejada relacionada ao uso de um medicamento (ICH, 1994) e vem sendo abordada de maneira semelhante aos EA na literatura. Uma tarefa complexa na detecção de RAM é determinar se o evento é decorrente do uso do medicamento ou da própria doença, ou seja, provar a relação causal entre a ocorrência e o medicamento administrado. Nesse sentido, Haerian *et al.* (2012) desenvolveram uma

abordagem de PLN para analisar se o medicamento é ou não responsável pela ocorrência. Como resultado, além da sensibilidade esperada, foi percebido que a cada hora investida utilizando a ferramenta, 20 horas foram economizadas na revisão manual de registros hospitalares.

Rever os potenciais acontecimentos adversos relacionados a medicamentos prescritos é uma tarefa clínica necessária e comum para os médicos (ARONSON *et al.*, 2002; ZAROWITZ *et al.*, 2005). Executar essa avaliação pode levar muito tempo, ainda mais se os pacientes estão tomando vários medicamentos (ELY *et al.*, 2005; RAMOS *et al.*, 2003). Embora a revisão manual de registros hospitalares seja comum em muitos estudos sobre EA, esse é um processo que utiliza muitos recursos, sendo caro e consumindo muito tempo (GOVINDAN *et al.*, 2010; KLOMPAS & YOKOE, 2009).

Um problema comum na detecção de RAM é que a mineração de grandes volumes de dados pode resultar em um grande número de possíveis pares de eventos adversos a medicamentos, que terão então de ser categorizados em grupos de eventos conhecidos e desconhecidos manualmente. Este é um processo demorado, mas que pode ser acelerado se uma base de dados de drogas relacionadas com um determinado fenótipo de EA estiver disponível (WANG, W. *et al.* 2011).

Kuhn *et al.* (2010) desenvolveram a fonte de dados, Side Effect Resource (SIDER), que tem apoiado diversos estudos sobre RAM. A base de dados disponibilizou a associação entre 888 medicamentos e 1450 efeitos colaterais. Essa base de dados é gratuita e disponível para estudos acadêmicos. Da mesma forma, Duke e Friedlin (2010) apresentam uma base de conhecimento e um serviço de apoio à decisão projetado para resolver as limitações das informações sobre EA a medicamentos. Primeiramente, os autores apresentam uma ferramenta de PLN usada para criar e manter essa base de conhecimento. Em seguida, é descrita a implementação e avaliação de desempenho do serviço de apoio à decisão que retorna dados de EA específicos após o recebimento de um documento de continuidade assistencial.

Já Gurulingappa *et al.* (2012), com o objetivo de facilitar a validação de estudos envolvendo mineração de textos e EA, desenvolveram um corpus textual sistematicamente anotado, com base no sistema online de busca e análise de literatura médica, Medline. Esse corpus foi definido como um padrão-ouro sobre medicamentos, efeitos adversos, dosagens e a relação entre eles, com o objetivo de

apoiar o desenvolvimento e validação de métodos para a extração automática de RAM.

A literatura médica também apresenta potencial para apoiar a detecção e análise das diversas reações a medicamentos e suas diversas associações a outros eventos, conforme demonstrado por Shang *et al.* (2014). O conhecimento extraído da literatura médica por meio de métodos de mineração de textos pode ser útil para a construção de padrões de RAM utilizáveis em outros estudos. Uma das vantagens da utilização do texto acadêmico como fonte de dados é o fato de os resultados publicados na literatura já terem sido avaliados por profissionais, tarefa fundamental da farmacovigilância.

Nesse sentido, Xu e Wang (2014) combinaram sinais da literatura e do FAERS (FDA Adverse Event Report System) para melhorar a vigilância de medicamentos. O sistema de relatório de eventos adversos FAERS tem sido usado em diversos estudos sobre mineração de textos. Por se tratar de uma base de dados pública, diversas pesquisas têm tirado proveito para avaliar e validar diferentes abordagens de mineração de textos, no entanto é necessário realizar a normalização dos dados para comparação com outras bases de dados, a qual pode ser apoiada pelo método proposto por Wang *et al.* (2014).

Também utilizando o FAERS, Ramesh *et al.*, (2014) tiveram como objetivo desenvolver um corpus de narrativas anotado e um anotador de entidades biomédicas nomeadas para identificar automaticamente informações de RAM relacionadas as narrativas FAERS. A melhor performance obteve um resultado de 0,73 de F-escore para a anotação de medicamentos, reações adversas e outras entidades nomeadas.

Já para extrair efeitos adversos de relatos de casos da literatura (MEDLINE) Gurulingappa *et al.* (2013) aplicaram um sistema de extração de relações linguísticas (Java Simple Relationship Extraction). Os autores tiveram como objetivo avaliar o impacto de sinais de EA detectados automaticamente a partir de texto e de dados de código aberto na previsão de alteração de rótulos de drogas. Como resultado, 76% de alteração de rótulos de drogas foram previstos automaticamente, concluindo que mudanças nas etiquetas das drogas podem ser previstas automaticamente usando técnicas de mineração de dados e texto.

A partir da extração de medicamentos e suas entidades relacionadas no texto, como doenças, sintomas, indicações, contraindicações, dosagem e reações adversas, é possível comparar a correspondência de registros hospitalares com as indicações

dos rótulos dos medicamentos de uma base de dados, como a base do Food and Drug Administration (FDA). Dessa forma, tendo como objetivo apoiar a detecção de RAM, Li *et al.* (2013) construíram um sistema para extração de medicamentos e suas entidades relacionadas do texto clínico.

Também com o objetivo de melhorar os métodos de detecção de eventos da farmacovigilância, Roitmann *et al.* (2014) apresentaram uma abordagem para analisar correlações entre EA e medicamentos por meio do agrupamento de EA, medicamentos e pacientes, buscando encontrar padrões interessantes para a farmacovigilância. Já Vilar *et al.* (2012) desenvolveram um modelo baseado em impressões digitais moleculares (MFBM) para reforçar a detecção de sinais de eventos adversos a medicamentos. Os autores utilizaram o sistema MedLEE para processar os dados narrativos não estruturados e extrair e normalizar entidades clínicas relevantes, tais como medicamentos, doenças, sintomas e informações temporais.

2.3.4.3 REAÇÕES ADVERSAS A VACINAS

Técnicas de mineração de textos e PLN também têm sido utilizadas na detecção de EA relacionados a vacinas, utilizando, por exemplo, a base de dados do sistema de eventos adversos a vacinas, Vaccine Adverse Event Reporting System (VAERS) (BAER, B. *et al.* 2016; BOTSIS *et al.* 2011; BOTSIS *et al.* 2012; BOTSIS *et al.* 2013; COURTOT *et al.*, 2014). Uma das reações adversas relacionada a vacinas pode ser a febre. A febre é uma elevação anormal de temperatura do corpo, geralmente em consequência de um processo patológico, sendo um dos EA mais comuns de vacinas (HUR *et al.*, 2012). Os mecanismos exatos das interações de genes associados à febre após vacina não são completamente compreendidos, o que dificulta a sua análise em textos clínicos.

Hur *et al.* (2012) tiveram como foco a detecção de febre após a imunização por vacinas, buscando compreender interações genéticas com base em publicações na literatura, extraídas do portal de publicações médicas PubMed. A pesquisa resultou em uma rede que consiste em 403 genes e 577 interações de genes relacionados a febre após vacinas. Já Hazlehurst *et al.* (2009), também buscando detectar possíveis EA após o uso de vacinas, modificaram o sistema Mediclass

(HAZLEHURST *et al.*, 2005), obtendo resultados positivos com o uso do processamento de linguagem natural em notas clínicas.

2.3.4.4 INTERAÇÕES MEDICAMENTOSAS

Novos medicamentos são testados interagindo com medicamentos existentes antes da aprovação do mercado usando métodos *in vivo* e *in vitro*. (ZHANG *et al.*, 2009). Devido ao grande número de maneiras em que as drogas podem interagir, pode ser inviável realizar testes para cada tipo de interação (KRONER, 2002). Muitas interações medicamentosas se manifestam após um certo período de exposição e pode demorar para interações raras ocorrerem (TRIARIDIS *et al.*, 2009). Portanto, o monitoramento pós comercialização é necessário para detectar interações imprevistas que ocorrem quando o medicamento está sendo utilizado pela população em geral.

Técnicas para detecção de EA derivados de interações medicamentosas têm utilizado dados de relatórios de EA do FDA (JIANG *et al.*, 2015; IYER *et al.*, 2014; SAMPATHKUMAR *et al.*, 2014; TATONETTI *et al.*, 2011). Como exemplo, é possível citar Tatonetti *et al.* (2011), que fizeram uma descoberta ao analisar a interação entre dois medicamentos, Paroxetina e Pravastatina no FAERS. Foram avaliados 239 pacientes retrospectivamente. Como resultado, foi constatado que os medicamentos possuem um efeito sinérgico de aumento na glicose no sangue.

Xu e Wang (2013) desenvolveram uma abordagem de aprendizagem de máquina *bootstrapping* semi-supervisionada, para extrair e ranquear de forma iterativa, pares de drogas e genes de acordo com a sua relevância para a farmacogenômica. A farmacogenômica desempenha um papel importante na identificação de respondedores e não-respondedores a medicamentos, evitando efeitos adversos e otimizando a dosagem do medicamento (WEINSHILBOUM e WANG, 2006; RODEN e TYNDALE, 2011).

2.4 CONSIDERAÇÕES FINAIS

Este trabalho buscou identificar, a partir de uma revisão sistemática da literatura, áreas da saúde e técnicas de mineração de textos utilizadas na prevenção

e detecção de EA. A estratégia utilizada foi a revisão sistemática da literatura que permitiu verificar as diferentes abordagens que vêm sendo apresentadas por pesquisadores.

As soluções de mineração de bases de dados textuais encontradas apontam predominantemente para a farmacovigilância e EA decorrentes do uso de medicamentos. É importante destacar o papel que as redes sociais vêm desempenhando como fonte de dados para a farmacovigilância, uma vez que a interação entre médicos e pacientes geralmente se limita aos períodos de internação e consultas. A vigilância das interações medicamentosas e de vacinas também vem sendo apoiada por métodos de mineração de textos. Rótulos de medicamentos e textos extraídos da literatura têm sido utilizados para apoiar a análise e descoberta de efeitos adversos antes mesmo da comercialização de medicamentos.

Das 54 pesquisas que fizeram parte deste estudo, 38 pesquisas estão relacionadas a medicamentos, incluindo pesquisas sobre vacinas, reações e efeitos adversos de interações medicamentosas. As demais 16 pesquisas abordam EA de maneira geral ou específica, como infecções nosocomiais e do sítio cirúrgico, quedas e tromboembolismo venoso. A maior parte das pesquisas utilizaram bases de dados públicas que permitem a continuação e a colaboração dos estudos entre diferentes pesquisadores.

Percebe-se como tendência a utilização de mais de um método ou técnica de mineração de textos na mesma solução, como a combinação de vários classificadores nos sistemas *ensemble*. Das 54 publicações analisadas, 32 utilizaram mais de um método de mineração de textos para detectar ou apoiar a identificação de EA. A maioria das publicações teve por objetivo apresentar novas propostas, sistemas, métodos ou modelos envolvendo mineração de textos para apoiar a vigilância de EA. No entanto, métodos tradicionais de MT como Suport Vector Machines (SVM), classificadores bayesianos e algoritmos de associação também são utilizados.

Quanto a análise da qualidade das publicações segundo o questionário do apêndice A, em média 85,84% foram atendidas totalmente, 11,44% atendidas parcialmente e 2,72% das questões não foram atendidas. Entre as questões que impactaram negativamente a qualidade das publicações está a não descrição do conjunto de dados utilizado: 35,19% das publicações não descreveram ou descreveram parcialmente os dados que constituíram a análise. Outro fator negativo foi a não descrição dos métodos estatísticos utilizados. Por fim, a não utilização de

dados públicos também é citada como aspecto negativo em 37% das publicações, em função da dificuldade de replicação dos estudos.

Como trabalhos futuros, sugere-se realizar a análise de múltiplos algoritmos de classificação textual em áreas pouco abordadas na literatura. Técnicas de classificação em documentos como prontuários, resultados de exames, prescrições de medicamentos e dados cirúrgicos apresentam potencial para utilização de mineração de dados textuais que podem indicar padrões para a melhoria da segurança do paciente nos processos assistenciais de saúde.

2.5 REFERÊNCIAS

- AGGARWAL, C.C., ZHAI, C.X. Mining Text Data. Springer Science + Business Media, LLC 2012.
- ALLEN, J. Natural language understanding. Redwood City, CA: Benjamin/Cummings Publishing Company, 1995.
- ARONSON, J.K., DERRY, S., LOKE, Y.K. Adverse drug reactions: keeping up to date. *Fundamental and Clinical Pharmacology*, 2002.
- BATES, D.W., BOYLE, D.L., VLIET, M. B. V. Relationship between medication errors and adverse drug events. *Journal of General Internal Medicine*, 1995.
- BATES, D.W., EVANS, R.S., MURFF, H., STETSON, P.D., PIZZIFERRI, L., HRIPCSAK, G. Detecting adverse events using information technology. *Journal American Medical Informatics Association*, 2003.
- BAUD, R.H., RASSINOX, A.M., WAGNER, J.C., LOVIS, C., JUGE, C., ALPAY, L.L., *et al.* Representing clinical narratives using conceptual graphs. *Methods of Information in Medicine*, 1995.
- BECKMANN, U., BOHRINGER, C., CARLESS, R., *et al.* Evaluation of two methods for quality improvement in intensive care: facilitated incident monitoring and retrospective medical chart review. *Critical Care Medicine*, 2003.
- BIEMANN, C., MEHLER, A. Text Mining, Theory and Applications of Natural Language Processing. Springer International Publishing Switzerland, 2014.
- CHAPMAN, W.W. Natural language processing for biosurveillance. In: Wagner MM, Moore AW, Aryel RM, eds. *Handbook of biosurveillance*. Burlington, MA: Elsevier Academic Press, 2006.
- CROFT, W. B., METZLER, D., STROHMA, T. Search Engines – Information Retrieval in Practice, Pearson Education, 2009.
- DEMNER-FUSHMAN, D.; CHAPMAN, W. W.; MCDONALD, C. J. What can natural language processing do for clinical decision support? *Journal of Biomedical Informatics*, 2009.
- ELY, J.W., OSHEROFF, J.A., CHAMBLISS, M.L., EBELL, M.H., ROSENBAUM, M.E. Answering Physicians' Clinical Questions: Obstacles and Potential Solutions. *Journal of the American Medical Informatics Association*, 2005.
- FAYYAD, U.; PIATETSKY-SHAPIRO, G; SMYTH, P. KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, 1996.
- FRIEDMAN, C., ELHADAD, N. Natural Language Processing in Health Care and Biomedicine. In: SHORTLIFFE, E.H., CIMINO, J.J. *Biomedical Informatics*. Springer-Verlag London. 2014.

GANZ, D.A., BAO, Y., SHEKELLE, P.G. Will my patient fall? Journal of American Medical Informatics Association. 2007.

GEERTS, W. H., BERGQVIST, D., PINEO, G. F. Prevention of venous thromboembolism: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines - 8th Edition, Chest 2008.

GOVINDAN, M., CITTERS, A. D. V., NELSON, E. C., KELLY-CUMMINGS, J., SURESH, G. Automated detection of harm in healthcare with information technology: a systematic review. Quality Safety Health Care, 2010.

GRIFFIN, F.A. RESAR, R.K. IHI Global Trigger Tool for Measuring Adverse Events (Second Edition). IHI Innovation Series white paper. Cambridge, Massachusetts: Institute for Healthcare Improvement, 2009.

HARPAZ, R., CALLAHAN, A., TAMANG, S., LOW, Y., ODGERS, D., FINLAYSON, S., JUNG, K., LEPENDU, P., SHAH, N. H. Text Mining for Adverse Drug Events: the Promise, Challenges, and State of the Art. Drug safety, 2014.

HEARST, M. A. Untangling text data mining. Proceedings 37th Annual meeting of the Association for Computational Linguistics. College Park, MD, 1999.

HIATT, H. H., BARNES, B. A., BRENNAN, T. A., LAIRD, N. M., LAWTHERS, A. G., LEAPE, L. L., LOCALIO, A. R., NEWHOUSE, J. P., PETERSON, L. M., THORPE, K. E. A study of medical injury and medical malpractice. An overview. New England Journal of Medicine, 1989.

HU, M., LIU, B. Mining opinion features in customer reviews. In: Proceedings of the National Conference on Artificial Intelligence. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 2004.

INTERNATIONAL CONFERENCE ON HARMONISATION. Clinical Safety Data Management: Definitions and Standards for Expedited Reporting E2a. International Conference on Harmonisation of Technical Requirements for Registration Of Pharmaceuticals For Human Use, 1994.

INSTITUTE OF MEDICINE. Crossing the quality chasm: a new health system for the 21st century. Washington, D.C.: National Academy Press, 2001.

INSTITUTE OF MEDICINE. Finding What Works in Health Care: Standards for Systematic Reviews. Washington, DC: The National Academies Press, 2011.

JHA, A. K., KUPERMAN, G. J., TEICH, J. M., LEAPE, L. L., SHEA, B., RITTENBERG, E., BURDICK, E., VANDERVLIT, M., SEGER, D., BATES, D. W. Identifying adverse drug events: development of a computer- based monitor and comparison with chart review and simulated voluntary report. Journal American Medical Informatics Association, 1998.

KLOMPAS, M., YOKOE, D.S. Automated surveillance of health care-associated infections. Clinical Infectious Diseases, 2009.

MENDONÇA, E.A., HAAS, J., SHAGINA, L., LARSON, E., FRIEDMAN, C. Extracting information on pneumonia in infants using natural language processing of radiology reports. *Journal of Biomedical Informatics*. 2005.

MEYSTRE S. M., SAVOVA, G.K., KIPPER-SCHULER, K.C., HURDLE, J.F. Extracting information from textual documents in the electronic health record: a review of recent research. *IMIA Yearbook of Medical Informatics*. 2008.

MITCHELL, T.M. *Machine learning*. Boston, MA: McGraw-Hill, 1997.

MOHER, D. LIBERATI, A., TETZLAFF, J., ALTMAN, D. G. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 6(7), 2009.

NABHAN, M., ELRAIYAH, T., BROWN, D., DILLING, J., LEBLANC, A., MONTORI, V., MORGENTHALER, T., NAESSENS, J., PROKOP, L., ROGER, V., SWENSEN, S., THOMPSON, R. L., MURAD, M. H. What is preventable harm in healthcare? A systematic review of definitions. *BMC Health Services Research*, 2012.

NEBEKER, J.R., BARACH, P., SAMORE, M.H. Clarifying adverse drug events: a clinician's guide to terminology, documentation, and reporting. *Annals of Internal Medicine*, 2004.

O'NEIL, A.C., PETERSEN, L.A., COOK, E.F., *et al.* Physician reporting compared with medical record review to identify adverse medical events. *Ann Intern Med*, 1993.

POLIKAR, R. Ensemble based systems in decision making. *IEEE Circuits and Systems Magazine*, 2006.

RADJENOVIĆ, D., HERIČKO, M., TORKAR, R., ŽIVKOVIČ., A. SOFTWARE FAULT PREDICTION METRICS: A SYSTEMATIC LITERATURE REVIEW. *Information and Software Technology*, 2013.

RAMOS, K., LINSCHIED, R., SCHAFFER, S. Real-time informationseeking behavior of residency physicians. *Family Medicine*, 2003.

RODEN, D. M., TYNDALE, R. F. Pharmacogenomics at the tipping point: challenges and opportunities. *Clinical Pharmacology e Therapeutics*, 2011.

SAGER, N., LYMAN, M., NHAN, N, T., TICK, L.J. Medical language processing: applications to patient data representation and automatic encoding. *Methods of Information in Medicine*, 1995.

SEBASTIANI, F. Machine learning in automated text categorization. *ACM Computing Surveys*, 2002.

SHAREK, P. J., CLASSEN, D. The incidence of adverse events and medical error in pediatrics. *Pediatric Clinics of North America*, 2006.

SIMPSON, M. S., FUSHMAN, D. D. Biomedical Text Mining: A Survey of Recent Progress. Springer International Publishing Switzerland. 2014. In: AGGARWAL, C.C., ZHAI, C.X. Mining Text Data. Springer Science+Business Media, 2012.

SPINKS, T., GANZ, P.A., SLEDGE, G. W. Delivering high-quality cancer care: The critical role of quality measurement. Healthcare, 2014.

SPYROPOULOS, A. C., HUSSEIN, M., LIN, J. Rates of venous thromboembolism occurrence in medical patients among the insured population. Thrombosis and Haemostasis, 2009.

WEINICK, R. M., BURNS, R. M., MEHROTRA, A. Many emergency department visits could be managed at urgent care centers and retail clinics. Health Affairs, 2010.

WEINSHILBOUM, R. M., WANG, L. Pharmacogenetics and pharmacogenomics: development, science, and translation. Annual Review of Genomics and Human Genetics, 2006.

WORLD HEALTH ORGANIZATION (WHO). The importance of pharmacovigilance-Safety Monitoring of Medicinal Products. Geneva: World Health Organization; 2002.

WORLD HEALTH ORGANIZATION (WHO). WHO draft guidelines for adverse event reporting and learning systems — from information to action. Geneva: World Health Organization, 2005.

ZAPT, D., REASON, J.T. Introduction to error handling. Applied Psychology, 1994.

ZAROWITZ, B.J., STEBELSKY, L.A., MUMA, B.K., ROMAIN, T.M., PETERSON, E.L. Reduction of high-risk polypharmacy drug combinations in patients in a managed care setting. Pharmacotherapy, 2005.

2.6 REFERÊNCIAS DA REVISÃO SISTEMÁTICA

ARAMAKI, E., MIURA, Y., TONOIKE, M., OHKUMA, T., MASUICHI, H., WAKI, K., OHE, K. Extraction of adverse drug effects from clinical records. Studies in Health Technology and Informatics, 2010.

BAER, B., NGUYEN, M., WOO, E. J., WINIECKI, S., SCOTT, J., MARTIN, D., BOTSIS, T., BALL, R. *et al.* Can natural language processing improve the efficiency of vaccine adverse event report review? Methods of Information in Medicine, 2016.

BENTON, A., UNGAR, L., HILL, S., HENNESSY, S., MAO, J., CHUNG, A., LEONARD, C. E., HOLMES, J. H. Identifying potential adverse effects using the web: A new approach to medical hypothesis generation. Journal of Biomedical Informatics, 2011.

BIAN, J., TOPALOGLU, U., YU, F. Towards large- scale twitter mining for drug- related adverse events. In Proceedings of the 2012 International Workshop on Smart Health and Wellbeing, 2012.

BOTSIS, T., NGUYEN, M. D., WOO, E. J., MARKATOU, M., BALL, R. Text mining for the Vaccine Adverse Event Reporting System: medical text classification using informative feature selection. Journal of the American Medical Informatics Association, 2011.

BOTSIS, T., BUTTOLPH, T., NGUYEN, M. D., WINIECKI, S., WOO, E. J., BALL, R. Vaccine adverse event text mining system for extracting features from vaccine safety reports. Journal of the American Medical Informatics Association, 2012.

BOTSIS, T., WOO, E. J., BALL, R. The contribution of the vaccine adverse event text mining system to the classification of possible Guillain-Barré syndrome reports. Applied clinical informatics, 2013.

CAMPILLO-GIMENEZ, B., GARCELON, N., JARNO, P., CHAPPLAIN, J. M., CUGGIA, M. Full- text automated detection of surgical site infections secondary to neurosurgery in Rennes, France. Studies in health technology and informatics, 2013.

CHEE, B. W., BERLIN, R., SCHATZ, B. Predicting adverse drug events from personal health messages. AMIA Annual Symposium proceedings, 2011.

CLARK, K., SHARMA, D., QIN, R., CHUTE, C. G., TAO, C. A use case study on late stent thrombosis for ontology-based temporal reasoning and analysis. Journal of biomedical semantics, 2014.

COURTOT, M., BRINKMAN, R., RUTTENBERG, A. The Logic of Surveillance Guidelines: An Analysis of Vaccine Adverse Event Reports from an Ontological Perspective. PLoS One, 2014.

BRUIJN, B., CHERRY, C., KIRITCHENKO, S., MARTIN, J., ZHU, X. Machine- learned solutions for three stages of clinical information extraction: the state of the art at i2b2 2010. Journal of the American Medical Informatics Association, 2011.

FREIFELD, C. C., BROWNSTEIN, J. S., MENONE, C. M., BAO, W., FILICE, R., KASSHOUT, T., DASGUPTA, N. Digital Drug Safety Surveillance: Monitoring Pharmaceutical Products in Twitter. Drug Safety, 2014

GENKIN, A., LEWIS, D. D., MADIGAN, D. Large-Scale Bayesian Logistic Regression for Text Categorization. Technometrics, 2007.

GURULINGAPPA, H., TOLDO, L., RAJPUT, A. M., KORS, J. A., TAWHEEL, A., TAYROUZ, Y. Automatic detection of adverse events to predict drug label changes using text and data mining techniques. Pharmacoepidemiology and Drug Safety, 2013.

GURULINGAPPA, H., RAJPUT, A. M., ROBERTS, A., FLUCK, J., HOFMANN-APITIUS, M. TOLDO, L. Development of a benchmark corpus to support the automatic extraction of drug-related adverse effects from medical case reports. Journal of Biomedical Informatics, 2012.

HAERIAN, K., VARN, D., VAIDYA, S., ENA, L., CHASE, H. S., FRIEDMAN, C. Detection of Pharmacovigilance- Related Adverse Events Using Electronic Health Records and Automated Methods. *Clinical Pharmacology & Therapeutics*, 2012.

HAZLEHURST, B., FROST, H. R., SITTING, D. F., STEVENS, V. J. MediClass: A System for Detecting and Classifying Encounter-based Clinical Events in Any Electronic Medical Record. *Journal of the American Medical Informatics Association*, 2005.

HAZLEHURST, B.; NALEWAY, A.; MULLOOLY, J. Detecting possible vaccine adverse events in clinical notes of the electronic medical record. *Vaccine*, 2009.

HONIGMAN, B., LEE, J., ROTHSCHILD, J., LIGHT, P., PHARM, D., PULLING, R. M., YU, T., BATES, D. W. Using computerized data to identify adverse drug events in outpatients. *Journal of the American Medical Informatics Association*, 2001.

HUR, J., ÖZGÜR, A., XIANG, Z., HE, Y. Identification of fever and vaccine-associated gene interaction networks using ontology-based literature mining. *Journal of Biomedical Semantics*, 2012.

IYER, S. V., HARPAZ, R., LEPENDU, P., BAUER-MEHREN, A., SHAH, N. H. Mining clinical text for signals of adverse drug-drug interactions. *Journal of the American Medical Informatics Association*, 2014.

DUKE, J. D.; FRIEDLIN, J. ADESSA: A Real-Time Decision Support Service for Delivery of Semantically Coded Adverse Drug Event Data. *AMIA Annual Symposium proceedings*, 2010.

HADZI-PURIC, J., GRMUSA, J. Automatic drug adverse reaction discovery from parenting websites using disproportionality methods. In *Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining*, 2012.

JIANG, G., LIU, H., SOLBRIG, H. R., CHUTE, C. G. Mining severe drug-drug interaction adverse events using Semantic Web technologies: a case study. *Bio Data Mining*, 2015.

KANG, N., AFZAL, Z., SINGH, B., MULLIGEN, E. M., KORS, J. A. Using an ensemble system to improve concept extraction from clinical records. *Journal of Biomedical Informatics*, 2012.

KUHN, M., CAMPILLOS, M., LETUNIC, I., JENSEN, L. J., BORK, P. A side effect resource to capture phenotypic effects of drugs. *Molecular Systems Biology*, 2010.

LEAMAN, R., WOJTULEWICZ, L., SULLIVAN, R., SKARIAH, A., YANG, J., GONZALEZ, G. Towards internet-age pharmacovigilance: extracting adverse drug reactions from user posts to health-related social networks. In *Proceedings of the 2010 workshop on biomedical natural language processing*, 2010.

LEPENDU, P., IYER, S. V., BAUER-MEHREN, A., HARPAZ, R., MORTENSEN, J. M., PODCHIYSKA, T., FERRIS, T. A., SHAH, N. H. Pharmacovigilance Using Clinical Notes. *Clinical Pharmacology & Therapeutics*, 2013.

MCCART, J. A., BERNDT, D. J., JARMAN, J., FINCH, D. K., LUTHER, S. L. Finding falls in ambulatory care clinical documents using statistical text mining. *Journal of the American Medical Informatics Association*, 2013.

MELTON, G. B.; HRIPCSAK, G. Automated Detection of Adverse Events Using Natural Language Processing of Discharge Summaries. *Journal of the American Medical Informatics Association*, 2005.

MICHELSON, J. D.; PARISEAU, J. S.; PAGANELLI, W. C. Assessing surgical site infection risk factors using electronic medical records and text mining. *American Journal of Infection Control*, 2014.

NIKFARJAM, A.; GONZALEZ, G. H. Pattern mining for extraction of mentions of Adverse Drug Reactions from user comments. *AMIA Annual Symposium proceedings*, 2011.

ONG, M.-S.; MAGRABI, F.; COIERA, E. Automated identification of extreme-risk events in clinical incident reports. *Journal of the American Medical Informatics Association*, 2012.

PERSEPHONE DOUPI, H. S., BRIAN BJØRN, ELLEN DEILKÅS, URBAN NYLÉN, HANSRUTBERG. Use of Global Trigger Tool in Detecting Adverse Events in Neurosurgery: Comparison of Manual and Text Mining Detection of Triggers. *Cognition, Technology & Work*, 2015.

RAMESH, B. P., BELKNAP, S. M., LI, Z., FRID, N., WEST, D. P., YU, H. Automatically Recognizing Medication and Adverse Event Information from Food and Drug Administration's Adverse Event Reporting System Narratives. *JMIR Medical Informatics*, 2014.

LI, Q., DELEGER, L., LINGREN, T., ZHAI, H., KAISER, M., STOUTENBOROUGH, L., JEGGA, A. G., COHEN, K. B., SOLTI, I. Mining FDA drug labels for medical conditions. *BMC Medical Informatics and Decision Making*, 2013.

LI, Q., MELTON, K., LINGREN, T., KIRKENDALL, E. S., HALL, E., ZHAI, H., NI, Y., KAISER, M., STOUTENBOROUGH, L., SOLTI, I. Phenotyping for patient safety: algorithm development for electronic health record based automated adverse event and medical error detection in neonatal intensive care. *Journal of the American Medical Informatics Association*, 2014.

ROCHEFORT, C. M., BUCKERIDGE, D. L., FORSTER, A. J. Accuracy of using automated methods for detecting adverse events from electronic health record data: a research protocol. *Implementation science*, 2015a.

ROCHEFORT, C. M., VERMA, A. D., EGUALE, T., LEE, T. C., BUCKERIDGE, D. L. A novel method of adverse event detection can accurately identify venous

thromboembolisms (VTEs) from narrative electronic health record data. *Journal of the American Medical Informatics Association*, 2015b.

ROITMANN, E., ERIKSSON, R., BRUNAK, S. Patient stratification and identification of adverse event correlations in the space of 1190 drug related adverse events. *Frontiers in Physiology*, 2014.

ROSS, M. K., LIN, K., TRUONG, K., KUMAR, A., CONWAY, M. Text Categorization of Heart, Lung, and Blood Studies in the Database of Genotypes and Phenotypes (dbGaP) Utilizing n-grams and Metadata Features. *Biomedical Informatics Insights*, 2013.

SAMPATHKUMAR, H., CHEN, X.-W., LUO, B. Mining Adverse Drug Reactions from online healthcare forums using Hidden Markov Model. *BMC medical informatics and decision making*, 2014.

SHANG, N., XU, H., RINDFLESCH, T. C., COHEN, T. Identifying plausible adverse drug reactions using knowledge extracted from the literature. *Journal of Biomedical Informatics*, 2010.

TAMANG, S., PATEL, M. I., BLAYNEY, D. W., KUZNETSOV, J., FINLAYSON, S. G., VETTETH, Y., SHAH, N. Detecting unplanned care from clinician notes in electronic health records. *Journal of oncology practice*, 2015.

TATONETTI, N. P., DENNY, J. C., MURPHY, S. N., FERNALD, G. H., KRISHNAN, G., CASTRO, V., YUE, P., TSAU, P. S., KOHANE, I. RODEN, D. M., ALTMAN, R. B. Detecting Drug Interactions From Adverse- Event Reports: Interaction Between Paroxetine and Pravastatin Increases Blood Glucose Levels. *Clinical Pharmacology & Therapeutics*, 2011.

TINOCO, A., EVANS, R. S., STAES, C. J., LLOYD, J. F., ROTHSCHILD, J. M., HAUG, P. J. Comparison of computerized surveillance and manual chart review for adverse events. *Journal of the American Medical Informatics Association*, 2011.

VILAR, S., HARPAZ, R., SANTANA, L., URIARTE, E., CAROL, C. Enhancing Adverse Drug Event Detection in Electronic Health Records Using Molecular Structure Similarity: Application to Pancreatitis. *PLoS ONE*, 2012.

WANG, L., JIANG, G., LI, D., LIU, H. Standardizing adverse drug event reporting data. *Journal of Biomedical Semantics*, 2014.

WANG, W., HAERIAN, K., SALMASIAN, H., HARPAZ, R., CHASE, H., FRIEDMAN, C. A drug- adverse event extraction algorithm to support pharmacovigilance knowledge mining from PubMed citations. *AMIA Annual Symposium proceedings*, 2011.

XU, R., WANG, Q. A semi-supervised approach to extract pharmacogenomics specific drug-gene pairs from biomedical literature for personalized medicine. *Journal of Biomedical Informatics*, 2013.

XU, R., WANG, Q. Large-scale combining signals from both biomedical literature and the FDA Adverse Event Reporting System (FAERS) to improve post-marketing drug safety signal detection. *BMC Bioinformatics*, 2014.

YANG, M., KIANG, M., SHANG, W. Filtering big data from social media – Building an early warning system for adverse drug reactions. *Journal of Biomedical Informatics*, 2015.

YELESWARAPU, S., RAO, A., JOSEPH, T., SAIPRADEEP, V. G., SRINIVASAN, R. A pipeline to extract drug- adverse event pairs from multiple data sources. *BMC Medical Informatics and Decision Making*, 2014.

1. APÊNDICE A - Questionário para avaliação da qualidade das publicações.

Questões	
1	Os objetivos (questões de pesquisa) são definidos com clareza?
2	O tamanho da amostra foi justificado?
3	Se o estudo envolve a avaliação do modelo de previsão / técnica, o modelo / técnica é claramente definidos?
4	As métricas utilizadas no estudo são relevantes para responder às questões de pesquisa?
5	As técnicas utilizadas no estudo são claramente definidas?
Método	
6	Os métodos de coleta de dados são descritos adequadamente?
Análise	
7	Os conjuntos de dados foram descritos adequadamente?
8	Os métodos estatísticos foram descritos?
9	Os métodos estatísticos se justificam?
10	O propósito da análise é claro?
11	Foram relatados efeitos colaterais?
Conclusão	
12	Todas as questões do estudo foram respondidas?
13	Os resultados apresentados são negativos?
14	O estudo discute como os resultados agregam valor para a literatura?
15	O estudo teve implicações práticas?
16	Os investigadores explicam as consequências de eventuais problemas com a validade / confiabilidade de suas medidas?
17	O estudo se repetível? Os conjuntos de dados usados são públicos?

2. APÊNDICE B – Métodos encontrados na literatura.

Área Geral	Métodos	Autores	Atendimento as questões de qualidade (%)
Eventos Adversos em Geral	Nomindex, Vector Space Model (VSM), UMLS	CAMPILLO-GIMENEZ <i>et al.</i> , 2013	88,24%
Eventos Adversos em Geral	Clinical Narrative Temporal Relation Ontology (CNTRO)	CLARK <i>et al.</i> , 2014	88,24%
Eventos Adversos em Geral	MetaMap, cTAKES, Medline, Support Vector Machine (SVM), Hidden Markov Model, OpenNLP	DE BRUIJN <i>et al.</i> , 2011	76,47%
Eventos Adversos em Geral	Logistic Regression (LR), Support Vector Machine (SVM), Ridge Logistic Regression (RLR)	GENKIN <i>et al.</i> , 2007	82,35%
Eventos Adversos em Geral	MediClass, UMLS, HL7's CDA	HAZLEHURST <i>et al.</i> , 2005	82,35%
Eventos Adversos em Geral	ACCCA (A combined clinical concept annotator), A Biomedical Named Entity Recognizer (ABNER), MetaMap, Lingpipe, JULIE Lab Named Entity Tagger (JNET), OpenNLP Chunker	KANG <i>et al.</i> , 2012	82,35%
Eventos Adversos em Geral	CCHMC's Automates Adverse Event Detection Program (AAEDP)	LI <i>et al.</i> , 2014	82,35%
Eventos Adversos em Geral	Logistic Regression (LR), Support Vector Machine (SVM), Cost-sensitive SVM (SVM-cost)	MCCART <i>et al.</i> , 2013	82,35%
Eventos Adversos em Geral	MedLEE	MELTON e HRIPCSAK, 2005	76,47%
Eventos Adversos em Geral	Eletronic Medical Record (EMR Epic Systems)	MICHELSON <i>et al.</i> , 2014	88,24%
Eventos Adversos em Geral	Naive Bayes (NB), Support Vector Machine (SVM)	ONG <i>et al.</i> , 2012	88,24%
Eventos Adversos em Geral	Global Trigger Tool (GTT), Text mining tool	PERSEPHONE DOUPI <i>et al.</i> , 2015	76,47%
Eventos Adversos em Geral	Automated AE Detection Algorithms	ROCHEFORT <i>et al.</i> , 2015a	70,59%
Eventos Adversos em Geral	Support Vector Machine (SVM)	ROCHEFORT <i>et al.</i> , 2015b	82,35%
Eventos Adversos em Geral	Naive Bayes (NB), Support Vector Machine (SVM), C4.5 <i>Decision Tree</i> , feature selection algorithm	ROSS <i>et al.</i> , 2013	82,35%
Eventos Adversos em Geral	Text mining algorithm	TAMANG <i>et al.</i> , 2015	70,59%
Eventos Adversos a Vacinas	VaeTM system	BAER <i>et al.</i> , 2016	88,24%
Eventos Adversos a Vacinas	Naive Bayes (NB), Maximum entropy (ME), <i>Decision Trees</i> (DT), Recursive Partitioning Classification Trees (RPCT), Boosted Trees (BT), <i>Weighted Support Vector Machine</i> (w-SVM), SVM for sparse data (s-SVM), Stochastic Boosting, Multivariate Adaptative Regression Splines (MARS), Regularized Discriminant Analysis (RDA), <i>Random Forests</i> (RF), Generalized Additive Model (GAM), <i>Weighted k-nearest Neighbors</i> (w-KNN).	BOTSIS <i>et al.</i> , 2011	88,24%
Eventos Adversos a Vacinas	VaeTM system	BOTSIS <i>et al.</i> , 2012	94,12%
Eventos Adversos a Vacinas	Vaccine adverse event Text Mining (VaeTM), UMLS	BOTSIS <i>et al.</i> , 2013	88,24%
Eventos Adversos a Vacinas	AERO (Adverse Events Reporting Ontology)	COURTOT <i>et al.</i> , 2014	94,12%
Eventos Adversos a Vacinas	Vaccine Safety Datalink (VSD), MediClass, UMLS	HAZLEHURST <i>et al.</i> , 2009	76,47%

Eventos Adversos a Vacinas	Support Vector Machine (SVM), CONDL (centrality and ontology-based network discovery using literature data)	HUR <i>et al.</i> , 2012	100%
Interações medicamentosas	Multi-Item Gamma Poisson Shrinker (MGPS), UMLS	IYER <i>et al.</i> , 2014	88,24%
Interações medicamentosas	Semantic Web-based, UMLS, SIDER, PharmGKB	JIANG <i>et al.</i> , 2015	94,12%
Interações medicamentosas	Latent Signal Detection Algorithm	TATONETTI <i>et al.</i> , 2011	70,59%
Interações medicamentosas	Computerized Surveillance System (CSS)	TINOCO <i>et al.</i> , 2011	58,82%
Reações Adversas a Medicamentos	Support Vector Machine (SVM)	ARAMAKI <i>et al.</i> , 2010	94,12%
Reações Adversas a Medicamentos	Association Rule Mining	BENTON <i>et al.</i> , 2011	88,24%
Reações Adversas a Medicamentos	Support Vector Machine (SVM), UMLS	BIAN <i>et al.</i> , 2012	94,12%
Reações Adversas a Medicamentos	Support Vector Machine (SVM)	CHEE <i>et al.</i> , 2011	88,24%
Reações Adversas a Medicamentos	Tree-based Dictionary-Matching Algorithm	FREIFELD <i>et al.</i> , 2014	100%
Reações Adversas a Medicamentos	ADESSA, SPLICER, UMLS	FRIEDLIN, 2010	76,47%
Reações Adversas a Medicamentos	Naive Bayes (NB), Maximum Entropy Classifier	GURULINGAPPA <i>et al.</i> , 2012	88,24%
Reações Adversas a Medicamentos	Peregrine, Java Simple Relationship Extraction (JSRE)	GURULINGAPPA <i>et al.</i> , 2013	88,24%
Reações Adversas a Medicamentos	Greedy Algorithm, UMLS	HADZI-PURIC e GRMUSA, 2013	88,24%
Reações Adversas a Medicamentos	Related Disease Identifier, UMLS, MedLEE	HAERIAN <i>et al.</i> , 2012	76,47%
Reações Adversas a Medicamentos	Micromedex M2D2	HONIGMAN <i>et al.</i> , 2001	82,35%
Reações Adversas a Medicamentos	Computer-Readable Side Effect Resource (SIDER), UMLS	KUHN <i>et al.</i> , 2010	70,59%
Reações Adversas a Medicamentos	Bayesian Confidence Neural Network (BCPNN), TPX framework	LEAMAN <i>et al.</i> , 2010	94,12%
Reações Adversas a Medicamentos	Association Rule Mining, UMLS	LEPENDU <i>et al.</i> , 2013	64,71%
Reações Adversas a Medicamentos	cTAKES, AutoMCEXtractor, UMLS	LI <i>et al.</i> , 2013	88,24%
Reações Adversas a Medicamentos	Association Rule Mining, UMLS	NIKFARJAM e GONZALEZ, 2011	94,12%
Reações Adversas a Medicamentos	Naive Bayes (NB), Support Vector Machine (SVM), Conditional Random Field (CRF), UMLS	POLEPALLI RAMESH <i>et al.</i> , 2014	94,12%
Reações Adversas a Medicamentos	Hierarchical Cluster (HC)	ROITMANN <i>et al.</i> , 2014	88,24%
Reações Adversas a Medicamentos	Hidden Markov Model (HMM) based Text Mining System, Viterbi decoding algorithm, Baum-Welch Algorithm	SAMPATHKUMAR <i>et al.</i> , 2014	94,12%
Reações Adversas a Medicamentos	Reflective Random Indexing (RRI), Predication-based Semantic Indexing (PSI), MetaMap, SemRep, UMLS	SHANG <i>et al.</i> , 2014	94,12%
Reações Adversas a Medicamentos	Disproportionality Analysis Algorithm, MedLEE, UMLS	VILAR <i>et al.</i> , 2012	100%
Reações Adversas a Medicamentos	Drug-ADE Detection Algorithm	WANG <i>et al.</i> , 2011	88,24%
Reações Adversas a Medicamentos	AERS-DM, MedEx, UMLS	WANG <i>et al.</i> , 2014	82,35%
Reações Adversas a Medicamentos	Ranking algorithm	XU e WANG, 2013	94,12%
Reações Adversas a Medicamentos	FAERS signals, FAERS signals augmented by pairs, Relative Reporting Ration (RRR), Proportional Reporting Ratio (PRR), Reporting	XU e WANG, 2014	94,12%

	Odds Ratio (ROR), phi Coefficient (PhiCorr), Information Component , UMLS		
Reações Adversas a Medicamentos	Support Vector Machine (SVM)	YANG <i>et al.</i> , 2015	94,12%
Reações Adversas a Medicamentos	Bayesian Confidence Neural Network (BCPNN)	YELESWARAPU <i>et al.</i> , 2014	94,12%

3 ARTIGO 2: Previsão de Infecções do Sítio Cirúrgico por meio da mineração de textos no Hospital de Clínicas de Porto Alegre

Resumo: O presente estudo foi realizado no Hospital de Clínicas de Porto Alegre (HCPA), Brasil, buscando verificar a hipótese de que o uso de métodos de mineração de texto pode ser uma alternativa viável para apoiar a previsão de Infecções em Sítio Cirúrgico (ISC). As ISC estão entre os EA mais comuns em pacientes hospitalizados e preveni-las é fundamental para a melhoria da segurança dos pacientes. Nesse sentido, este estudo teve por objetivo avaliar o desempenho de métodos de mineração de textos (MT) para prever ISC. Foram analisadas 15.479 descrições de cirurgias, nas quais foram aplicadas técnicas de pré-processamento de texto e métodos de classificação AdaBoost, DecisionTree, LinearSVC, Logistic Regression, Multinomial Naive Bayes, Nearest Centroid, Random Forest, Stochastic Gradient Descent, Support Vector Classification (SVC). O método Stochastic Gradient Descent apresentou o melhor desempenho pela métrica ROC-AUC (Receiver Operating Characteristic – Area Under the Curve), 79,7%, já o método Decision Tree apresentou o pior desempenho, 68,1%. Os métodos de MT podem ser usados para apoiar a previsão de infecções direcionando ações de vigilância de infecções no período pós-operatório.

Palavras-chave: Previsão de Eventos Adversos; Infecção do Sítio Cirúrgico; Mineração de Texto.

3.1 INTRODUÇÃO

Infecções em Sítio Cirúrgico (ISC) estão entre as principais infecções relacionadas à assistência à saúde no Brasil (ANVISA, 2013). As ISC também estão entre os Eventos Adversos (EA) mais comuns em pacientes hospitalizados, causando substancial aumento na mortalidade, taxas de reinternação e custos assistenciais (WATCHER, 2013; STONE *et al.*, 2009).

Os EA em pacientes cirúrgicos abrangem eventos comuns a todos os pacientes hospitalizados, como quedas, erros de diagnóstico, pneumonia nosocomial, entre outros, mas, além disso, abrangem também EA específicos de natureza técnica,

como as infecções em ferimentos e os sangramentos (TOWNSEND *et al.*, 2014). Nesse contexto, a Comissão de Controle de Infecção Hospitalar (CCIH) é o órgão encarregado por informar, vigiar, pesquisar, prevenir e controlar as infecções hospitalares (ANVISA, 2017).

No HCPA, a CCIH executa ações buscando a melhoria da segurança do paciente, incluindo pesquisas entre as quais o presente estudo se encontra. São monitorados processos que apresentam potencial impacto sistêmico, realizada a vigilância das infecções, instituídas medidas de precaução, auditoria da prescrição e elaboração de políticas de uso antimicrobianos da instituição (HCPA, 2017).

Identificar e medir EA é fundamental nos processos de saúde, formando uma base para a segurança do paciente (CLASSEN *et al.*, 2011). Métodos tradicionais para a prevenção e detecção de infecções podem utilizar recursos de maneira intensiva e demorada. Nos últimos anos, no entanto, a coleta e armazenamento de informações de saúde em formatos eletrônicos tornaram-se onipresentes, e vários bancos de dados eletrônicos do ambiente hospitalar, podem ser usados para realizar vigilância contínua de infecções (FREEMAN *et al.*, 2013).

Com o aumento da quantidade de texto disponível em diferentes sistemas, tem crescido a necessidade de utilização de métodos que possam aprender padrões interessantes nesses dados (AGGARWAL e ZHAI, 2012) e apoiem a tomada de decisão para melhoria da segurança do paciente. Nesse sentido, técnicas como o Processamento de Linguagem Natural (PLN), Mineração de Textos (MT), Aprendizagem de Máquina (AM), entre outras, vêm sendo utilizadas na área da saúde (KANG *et al.*, 2012; MEYSTRE *et al.*, 2008; DEMNER-FUSHMAN *et al.*, 2009).

O presente estudo tem por hipótese que métodos de MT podem ser usados para apoiar a tomada de decisões quanto a prevenção de ISC, prevendo a sua ocorrência a partir do texto de descrições de cirurgias. Dessa forma, o presente estudo tem por objetivo avaliar o desempenho de métodos de mineração de textos na previsão de ISC.

Este artigo está organizado em cinco seções além desta introdução. A seção 3.2 aborda o referencial teórico da pesquisa e as seções 3.3 e 3.4 apresentam a os procedimentos metodológicos e o processo de MT utilizado, respectivamente. A seção 3.5 destina-se às considerações finais.

3.2 REFERENCIAL TEÓRICO

A ocorrência de eventos adversos pode acarretar no aumento na morbidade, na mortalidade, no tempo de tratamento dos pacientes e nos custos assistenciais, além de repercutir em outros campos da vida social e econômica da população (ANVISA, 2013). Cerca de uma em 30 cirurgias “limpas” sofrerá complicações por uma ISC. A taxa é significativamente maior em cirurgias “sujas”, de emergência ou em procedimentos prolongados e para pacientes com comorbidades clínicas (WATCHER, 2013).

Existem estratégias para prevenir infecções em local cirúrgico, como: utilização de antibióticos profiláticos, utilização de cortadores para remoção de pelos antes da cirurgia, manutenção da normotermia após a cirurgia (YOKOE *et al.*, 2008; RANJI *et al.*, 2007; MELLING *et al.*, 2001), entre outras, que geralmente são documentadas nos registros eletrônicos do paciente. No contexto cirúrgico, as infecções são avaliadas de acordo com o seu potencial de contaminação a partir da seguinte classificação (GARNER *et al.*, 1988):

- a) Cirurgias limpas: sítio cirúrgico sem sinais de inflamação, sem contato com trato respiratório, alimentar, genital e urinário;
- b) Cirurgias potencialmente contaminadas: sítio cirúrgico envolvendo tratos respiratório, genital, gastrintestinal ou urinário em condições controladas e sem contaminação acidental;
- c) Cirurgias contaminadas: feridas abertas acidentalmente ou cirurgias com quebra importante de técnica asséptica ou grande contaminação do trato gastrintestinal. Cirurgias que entram no trato urinário com urina infecciosa ou trato biliar com bile infectada ou cirurgias onde é achado tecido inflamatório agudo não purulento;
- d) Cirurgias infectadas: lesões traumáticas antigas com tecido desvitalizado, corpo estranho, contaminação fecal, quando há perfuração inesperada de víscera.

A infecção hospitalar é um processo infeccioso decorrente de microorganismos presentes no ambiente hospitalar e corresponde a cerca de 10% de todas as internações hospitalares. Próximo de 30% das infecções hospitalares são evitáveis com medidas preventivas. Já a ISC é determinada por uma combinação de fatores: quantidade e tipo de contaminação, técnica cirúrgica e anestésica empregada

e resistência do hospedeiro. Essas infecções habitualmente são do local cirúrgico e contribuem para o aumento das taxas de mortalidade e tempo de permanência hospitalar (ROCHA, 2008). O potencial de contaminação, duração da cirurgia e a classificação ASA (American Society of Anesthesiologists), constituem o Índice de Risco de Infecção Cirúrgica (IRIC), que tem ajudado a prever ISC.

A vigilância em saúde tem sido descrita como um componente essencial de um programa eficaz de prevenção e controle das infecções, demonstrando conduzir a uma diminuição das taxas de infecção (HALEY *et al.*, 1985; BROWN *et al.*, 2006, CURRAN *et al.*, 2000). A busca passiva de infecções geralmente ocorre através de notificações espontâneas dos profissionais de saúde, porém, a maior parte das infecções pode não ser notificada. Por outro lado, a vigilância e a busca ativa de infecções, a partir de revisões pontuais, pode ajudar a identificar questões que necessitam de mais investigação, no entanto podem utilizar recursos de maneira intensiva e demorada além de não fornecer informação em tempo real, uma vez que o processo de coleta de dados, análise e *feedback* é extenso (FREEMAN *et al.*, 2013).

Para apoiar a vigilância em saúde, técnicas de Mineração de Dados (MD) têm sido utilizadas na literatura (ARAMAKI *et al.*, 2010; MCCART *et al.*, 2013; SILVA *et al.*, 2008). A MD é parte do processo de Descoberta de Conhecimento em Bases de Dados (DCBD), processo este que busca extrair conhecimento útil como padrões válidos e compreensíveis de bases de dados (FAYYAD, 1996). Han *et al.* (2006) afirmam que muitas vezes a MD de dados é utilizada por si só para descrever o processo de DCBD.

A mineração de dados utiliza como base algoritmos de Aprendizagem de Máquina (AM), os quais podem ser divididos em duas categorias principais: aprendizagem supervisionada e aprendizagem não supervisionada (HAN *et al.*, 2006). Na aprendizagem supervisionada as classes dos registros - infectadas e não infectadas - são conhecidas e também fornecidas como treino para que diferentes algoritmos possam extrair padrões que diferenciem os registros de uma classe ou de outra. Já na aprendizagem não supervisionada nenhum conjunto de treinamento é fornecido onde geralmente a tarefa dos algoritmos é realizar agrupamentos por similaridade, por exemplo (ZAFARANI *et al.*, 2014).

A MD aplicada a dados textuais não estruturados é chamada de MT, também tendo o princípio de extrair dos dados padrões não triviais, úteis e de interesse do usuário (TAN, 1999). Da mesma forma que a MD, a MT pode utilizar algoritmos de

AM, sendo possível obter previsões e classificações de novos registros com base no treinamento de registros existentes (AGGARWAL e ZHAI, 2012).

No contexto cirúrgico, a MT pode ser aplicada no texto inserido durante ou logo após a realização de uma cirurgia para prever a ocorrência de infecções. Neste caso, o método de MT deverá ter sido treinado em diversos registros nos quais infecções ocorreram ou não, e terá a tarefa de classificar se o texto diz respeito a ocorrência ou não de infecções.

O conhecimento exato das taxas de ISC pode ser usado como parte do mecanismo de feedback para diminuir a incidência futura de infecção no local cirúrgico (MICHELSON, 2005), além de ser usado no treinamento dos algoritmos de AM na MT. A identificação de ISC envolve interpretação de achados clínicos e laboratoriais (MANGRAM *et al.*, 1999). Todavia, pacientes cirúrgicos podem ficar pouco tempo no hospital após a cirurgia. Dessa forma, a identificação dos pacientes que desenvolvem infecção no local cirúrgico após a alta será o maior desafio nos métodos de vigilância de infecções (FREEMAN *et al.*, 2013).

É possível perceber, por meio da literatura, que a utilização de métodos de MT do processo de DCBD pode e vem trazendo benefícios para apoiar a análise de eventos que causam danos aos pacientes. Entretanto, se faz necessário o desenvolvimento de pesquisas que verifiquem a aplicabilidade desses métodos na previsão de ISC, propondo melhorias aos serviços de controle de infecções hospitalares no contexto dos hospitais brasileiros (DALTOE *et al.*, 2014).

3.3 PROCEDIMENTOS METODOLÓGICOS

Este estudo tem natureza aplicada e busca gerar soluções para a previsão de ISC no período pós-operatório. Este estudo também tem caráter descritivo, pois visa apresentar como métodos de MT podem ser aplicados para apoiar a previsão de ISC que ocorrerão em até 30 (trinta) dias depois da cirurgia.

Foram aplicados métodos de MT para avaliar registros cirúrgicos de um hospital brasileiro com a intenção de prever ISC. Foram utilizados dados disponibilizados pelo hospital para comparar técnicas de pré-processamento e classificação de texto. Trata-se de um estudo retrospectivo que utiliza métodos MT para prever infecções com base na descrição textual da cirurgia. O estudo foi realizado

em quatro etapas que vão desde a seleção dos dados do sistema informatizado do hospital até a avaliação de desempenho dos métodos de classificação, conforme Figura 3.

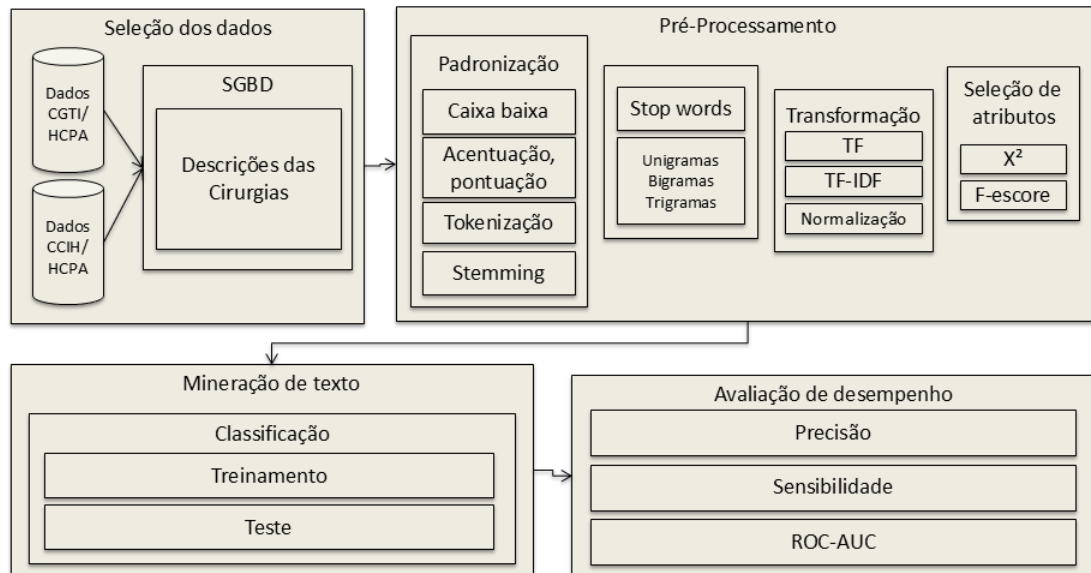


Figura 3. Procedimentos Metodológicos

Para gerenciar os dados do estudo, foi utilizado o Sistema Gerenciador de Banco de Dados (SGBD) PostgreSQL, versão 9.6 (POSTGRESQL, 2017). Já para o pré-processamento, aplicação de métodos de MT e avaliação de desempenho foi utilizada a linguagem de programação de alto nível, Python versão 3.5 (PYTHON, 2017). O Python inclui bibliotecas para apoiar o processamento dos dados, como Natural Language Toolkit (NLTK) (BIRD, *et al.* 2009) e Scikit-learn (PEDREGOSA, *et al.*, 2011), que foram utilizadas neste estudo.

A etapa de seleção dos dados teve por objetivo recuperar e integrar os dados da CCIH e das descrições das cirurgias do sistema informatizado do HCPA, preparando-os para o pré-processamento. A comissão de controle de infecções do HCPA, já identifica infecções por meio da revisão retrospectiva dos prontuários, possibilitando saber quais cirurgias tiveram ISC ou não. Os dados históricos já analisados no HCPA possibilitaram o treinamento de algoritmos de classificação, bem como a avaliação do desempenho da utilização de técnicas de MT para prever infecções após cirurgias.

O pré-processamento buscou preparar e estruturar o texto da descrição cirúrgica para as análises quantitativas dos métodos de classificação. Nesta etapa foi

realizada a padronização do texto, que inclui transformação do texto para caixa baixa, remoção de acentuação e pontuação, tokenização, *stemming*, remoção de *stop words*, e criação de unigramas, bigramas e trigramas, bem como os processos de transformação e seleção de atributos, que serão descritos detalhadamente nas seções a seguir.

Já na etapa da mineração de texto, foram aplicados os métodos de classificação AdaBoost, Decision Tree, Linear Support Vector Classification (LinearSVC), Logistic Regression, MultinomialNB, Nearest Centroid, Random Forest, Stochastic Gradient Descent (SGD), Support Vector Classification (SVC), que a partir do treinamento e teste com validação cruzada, identificaram cirurgias como positivas ou negativas para ISC no período pós-operatório. A seleção de parâmetros e estratégia de pré-processamento ocorreu por meio da busca aleatória em uma grade de parâmetros, com o objetivo de encontrar a melhor combinação de parâmetros para cada método (BERGSTRA e BENGIO, 2012).

Os métodos de classificação foram selecionados pela sua disponibilidade na biblioteca Scikit-learn (PEDREGOSA *et al.*, 2011) e utilização em estudos anteriores: (i) Support Vector Machines (SVM) (ARAMAKI *et al.*, 2010; BIAN *et al.*, 2012; BOTSIS *et al.*, 2011; BRUIJN *et al.*, 2011; CHEE *et al.*, 2011; HUR *et al.*, 2012; MCCART *et al.*, 2013; GENKIN *et al.*, 2007; ONG *et al.*, 2012; RAMESH *et al.*, 2014; ROCHEFORT *et al.*, 2015; ROSS *et al.*, 2013; YANG *et al.*, 2015); (ii) Logistic Regression (MCCART *et al.*, 2013; GENKIN *et al.*, 2007); (iii) Naive Bayes (BOTSIS *et al.*, 2011; GURULINGAPPA *et al.*, 2012; ONG *et al.*, 2012; RAMESH *et al.*, 2014; ROSS *et al.*, 2013); (iv) Decision Trees (BOTSIS *et al.*, 2011; FREIFELD *et al.*, 2014; ROSS *et al.*, 2013); (v) Boosted Trees, Random Forest, Nearest Neighbors (BOTSIS *et al.*, 2011).

Por fim, a avaliação do desempenho dos classificadores foi feita com base na sua Precisão, Sensibilidade e a Área Sob a Curva das Características Operacionais do Receptor (ROC-AUC). Foram mantidos na análise apenas os métodos que atingiram mais de 60% de ROC-AUC após a otimização de parâmetros.

3.4 MINERAÇÃO DE TEXTOS

3.4.1 SELEÇÃO DOS DADOS

O HCPA utiliza o sistema informatizado AGH (Aplicativos de Gestão Hospitalar) no qual todo o texto clínico inserido por médicos, enfermeiros e acadêmicos fica centralizado, viabilizando a coleta e posterior aplicação de métodos de MT. O texto cirúrgico foi disponibilizado pelo HCPA em planilhas do Microsoft Office Excel (.xlsx), onde cada linha corresponde a uma cirurgia. Esses registros foram então inseridos no Postgres, onde foi criada uma tabela para inserção dos dados consolidados, excluindo registros vazios ou com inconsistências. A tabela analisada foi formada por duas colunas, uma contendo o texto com a descrição da cirurgia e outra com um rótulo binário - sendo 1 para ocorrência de infecção e 0 para a não ocorrência de infecção - utilizado para classificar os registros.

O texto coletado de cada cirurgia possui a descrição técnica da cirurgia, mas também possui informações como: nome, sexo e idade do paciente, além da escala ASA, entre outros. A identificação de ISC trata-se da análise de um conjunto de dados com classes desbalanceadas, onde inicialmente, 98,8% dos registros possuem rótulo 0 (negativo para ocorrência de infecção) e apenas 1,2% possuem o rótulo 1 (positivo para ocorrência de infecção). Os rótulos dos registros infectados foram fornecidos pela CCIH, que acompanha mensalmente o número de infecções pós-operatórias no HCPA.

Foram excluídos da análise registros em que o registro não possui texto na descrição cirúrgica. Da mesma forma, foram excluídos os registros em que a ISC foi constatada com mais de 30 (trinta) dias após a cirurgia. Também foram excluídos da análise os registros de pacientes que realizaram mais de uma cirurgia no mesmo dia, das quais ao menos uma foi infectada.

3.4.2 PRÉ-PROCESSAMENTO

O pré-processamento teve por objetivo preparar o texto para a aplicação dos métodos de classificação. Apenas dados do campo textual foram tratados nessa etapa, nos quais, inicialmente, as letras maiúsculas foram substituídas por minúsculas

e a acentuação das palavras foi removida. A seguir, antes de um processamento mais sofisticado, o fluxo de palavras deve ser dividido em partes menores que tenham significado por meio de um processo chamado tokenização (FELDMANN e SANGER, 2006). Nesse processo os documentos podem ser divididos em capítulos, seções, parágrafos, frases, palavras e até mesmo sílabas ou fonemas. A abordagem mais comum utilizada em MT envolve quebrar o texto em palavras, sendo esta a abordagem utilizada neste estudo.

Após a tokenização, utilizando a biblioteca NLTK em português, foi possível realizar a remoção das *stop words*, que, de maneira geral, são palavras com alta frequência e que normalmente não contribuem para a semântica dos registros; por exemplo, a, o, de, da, entre outras (FELDMANN e SANGER, 2006). Após a remoção de *stop words* as palavras passaram por um processo de normalização morfológica, também chamada de *stemming*.

O processo de *stemming* busca remover afixos morfológicos das palavras deixando apenas o seu radical, com o objetivo de diminuir a dimensionalidade do texto. Por exemplo, as palavras infecção, infectado e infeccioso, são todas reduzidas para infec. Para realizar essa tarefa de *stemming* foi utilizado o algoritmo RSLPStemmer (ORENGO e HUYCK, 2001), disponível no NLTK.

Após as etapas de padronização, tokenização, remoção de *stop words* e *stemming*, cada registro se tornou uma lista de *tokens* seguidos pelo rótulo da sua classe, 0 ou 1. A partir disso, foi realizado o processo de definição de unigramas, bigramas ou trigramas, pois cada um desses veio a ser um atributo analisado. Um *n*-grama é uma subsequência de *n-tokens*; dessa forma, um unigrama corresponde a um único *token*, um bigrama a dois *tokens*, por exemplo uma palavra seguida de outra, e um trigrama corresponde a três *tokens*, três palavras consecutivas (PERKINS, 2014).

Na etapa seguinte, foi realizada a conversão do texto em uma matriz esparsa de contagem de termos (*tokens*), que possibilita a criação de análises quantitativas (SCIPY, 2017). Ao agregar a matriz esparsa a partir da frequência de termos, é possível representá-la como uma lista de documentos (registros), palavras e a sua frequência no documento, conforme Figura 4.

Registro	Palavra	Frequência
1	FO	2
1	secreção	2
1	purulenta	2
.	.	.
.	.	.
.	.	.
12435	dor	1

Figura 4. Matriz de termos

A Frequência do Termo (TF) no documento pode ser definida por $tf_{t,d}$, em que t e d denotam termo e documento respectivamente. Nesta visão de um documento, conhecido na literatura como o saco de palavras, a ordem exata dos termos em um documento é ignorada, mas o número de ocorrências de cada termo é mantido (MANNING, *et al.*, 2009). Ao analisar grandes volumes de documentos, termos interessantes, porém raros, podem permanecer ocultos e ter pouca influência no processo de classificação. Para combater esse problema a Frequência Inversa de Documentos também foi utilizada (IDF) (SKLEARN, 2017a), conforme Equação 1:

(1)

$$idf(t) = \log \frac{1 + n_d}{1 + df(d,t)}$$

Onde n_d é o número total de documentos, e $df(d,t)$ é o número de documentos em que o termo t aparece. O número 1 foi adicionado ao numerador e denominador como se um documento extra fosse visto contendo todo termo pelo menos uma vez, com o objetivo de evitar divisões por zero. Dessa forma, foi obtido o peso TF-IDF, onde TF é multiplicado por IDF, conforme descrito na Equação 2:

(2)

$$tf - idf(t,d) = tf(t,d) idf(t)$$

Para finalizar, o valor de TF-IDF para cada termo na matriz foi normalizado a partir da norma euclidiana unitária (SKLEARN, 2017a), conforme Equação 3:

(3)

$$v_{norm} = \frac{v}{\|v\|_p} = \frac{v}{\sqrt{v_1^p + v_2^p + \dots + v_n^p}}$$

Onde v_{norm} é o vetor normalizado, v é o vetor a ser normalizado e $\|v\|_p$ a norma utilizada, sendo $p = 1$ para ℓ_1norm e $p = 2$ para ℓ_2norm (YI e FALOUTSOS, 2000). Um alto valor TF-IDF é alcançado quando há uma alta frequência do termo (tf) em um dado documento e uma baixa frequência do termo no conjunto total de documentos, sendo possível perceber a importância de um termo em um determinado documento a partir de um alto valor TF-IDF.

3.4.3 SELEÇÃO DE ATRIBUTOS

A seleção de atributos no texto cirúrgico buscou selecionar atributos discriminatórios para determinada classe a partir de testes estatísticos. No que diz respeito a previsão de ISC, a seleção de atributos pode ser usada para descobrir e analisar fatores do processo cirúrgico que influenciam na ocorrência de infecções.

A seleção de atributos busca determinar quais termos são mais relevantes para o processo de classificação, dado que alguns termos são mais prováveis de serem relacionados a determinadas classes. No que diz respeito a classificação de textos, a seleção de atributos é especialmente importante devido à alta dimensionalidade de termos e a existência de ruídos (AGGARWAL e ZHAI, 2012).

Neste estudo, a seleção de atributos foi feita com base nas métricas qui-quadrado e F-escore, pela sua adequação ao problema e utilização na seleção de atributos em estudos anteriores (MICHELSON *et al.*, 2014; LUCINI *et al.*, 2017). Foram selecionados atributos com base em um percentil dos maiores valores, sendo que o percentil e a métrica utilizada variaram para cada método de MT analisado.

Na estatística, o teste qui-quadrado é aplicado para testar a independência de dois eventos, onde dois eventos A e B são definidos como independentes se $P(AB) = P(A)P(B)$, ou, de forma equivalente, $P(A | B) = P(A)$ e $P(B | A) = P(B)$, conforme Equação 4. Na seleção de atributos, os dois eventos são a ocorrência do termo e a ocorrência da classe (MANNING *et al.*, 2009):

(4)

$$X^2(D, t, c) = \sum_{e_t \in \{0,1\}} \sum_{e_c \in \{0,1\}} \frac{(N_{e_t e_c} - E_{e_t e_c})^2}{E_{e_t e_c}}$$

Onde $e_t = 1$ (o documento contém o termo t) e $e_t = 0$ (o documento não contém t), e $e_c = 1$ (o documento está na classe c) e $e_c = 0$ (o documento não está na classe c). Sendo N a frequência observada em D e E a frequência esperada. Dessa forma, valores superiores indicam maior dependência entre o termo e a classe (SEBASTIANI, 2002). Já a medida F -escore é definida pela Equação 5 (GUYON *et al.*, 2006):

(5)

$$F(i) = \frac{(\bar{x}_i^{(pos)} - \bar{x}_i)^2 + (\bar{x}_i^{(neg)} - \bar{x}_i)^2}{\frac{1}{n_{pos} - 1} \sum_{k=1}^{n_{pos}} (x_{k,i}^{(pos)} - \bar{x}_i^{(pos)})^2 + \frac{1}{n_{neg} - 1} \sum_{k=1}^{n_{neg}} (x_{k,i}^{(neg)} - \bar{x}_i^{(neg)})^2}$$

Onde \bar{x}_i , $\bar{x}_i^{(pos)}$, $\bar{x}_i^{(neg)}$ são a média da i -ésima característica dos conjuntos de dados completo, positivo e negativo, respectivamente, $x_{k,i}^{(pos)}$ é a i -ésima característica da k -ésima instância positiva (possui infecção), e $x_{k,i}^{(neg)}$ é a i -ésima característica da k -ésima instância negativa (não possui infecção). Já n_{pos} e n_{neg} são o número de casos positivos e negativos de infecção, respectivamente. Dessa forma, o numerador indica a discriminação entre os conjuntos positivo e negativo, e o denominador indica o que está dentro de cada um dos dois conjuntos (GUYON *et al.*, 2006). Quanto maior for a pontuação F , mais provável é que esta característica seja mais discriminativa.

3.4.4 CLASSIFICAÇÃO

Neste trabalho foram analisados vários métodos de classificação supervisionada. Nesse tipo de aprendizagem, os valores de atributo da classe para o conjunto de dados são conhecidos antes de executar o algoritmo. Esses dados são chamados dados rotulados ou dados de treinamento. Nesse modelo, os registros do conjunto de dados são tuplas no formato (x, y) , onde x é um vetor de termos tratados

no pré-processamento e y é o atributo de classe, 0 ou 1 para a ocorrência de infecção. A aprendizagem supervisionada constrói um modelo que mapeia x para y . Dessa forma, a tarefa de treinamento é encontrar um mapeamento $m()$ tal que $m(x) = y$ (ZAFARANI *et al.*, 2014). Para avaliar o modelo construído no treinamento, um conjunto de dados no formato $(x, ?)$ é utilizado para teste no qual o algoritmo deve atribuir 0 ou 1 com base no modelo aprendido no treinamento e uma posterior avaliação é realizada.

O presente estudo apresenta um desequilíbrio na distribuição das classes alvo: 98,79% para amostras negativas e 1,21% para amostras positivas para ocorrência de ISC. Para garantir que o modelo aprendido por cada algoritmo retivesse maior aproveitamento sobre a classe minoritária (positiva), foi utilizado o método de validação cruzada estratificada. Na validação cruzada, o conjunto de dados é dividido em k partições aproximadamente iguais, e cada uma por sua vez é usada para testes e o restante $(k - 1)$ é usada para treinamento. Nesse estudo, k foi definido igual a 10 (SKLEARN, 2017b).

Ao estratificar o treinamento em cada uma das partições na validação cruzada, os dados são divididos aleatoriamente de maneira que cada classe é representada na partição nas mesmas proporções que no conjunto de dados completo (98,4% e 1,6%). Dessa forma o processo de aprendizagem é executado um total de 10 vezes em diferentes conjuntos de treinamento, e por fim calcula-se a média dos 10 escores para obter uma estimativa global (WITTEN *et al.*, 2011).

Cada método de classificação possui parâmetros que não são diretamente aprendidos pelo classificador; por exemplo, C , kernel e gama para SVM (SKLEARN, 2017c). A definição desses parâmetros foi realizada por meio da busca aleatória em uma grade de parâmetros (BERGSTRA e BENGIO, 2012). Essa busca teve como objetivo encontrar a melhor combinação de diversos parâmetros para maximizar uma determinada métrica, que neste estudo foi a Área Sob a Curva das Características Operacionais do Receptor ponderada pelo balanceamento das classes (DONG *et al.*, 2012). A busca em grade também foi utilizada para selecionar o percentil e métrica (X^2 e F-escore) para seleção de atributos e norma para TF-IDF (l1 ou l2).

3.4.5 AVALIAÇÃO DO DESEMPENHO

A avaliação do desempenho dos classificadores foi feita a partir das métricas Precisão, Sensibilidade (*Recall*) e ROC-AUC. Adicionalmente matrizes de confusão foram analisadas. Essas métricas foram selecionadas pela sua adequação ao problema de classificação em conjuntos de dados desbalanceados e pela sua utilização em estudos anteriores (ARAMAKI *et al.*, 2010; BIAN *et al.*, 2012; BOTSIS *et al.*, 2011; BRUIJN *et al.*, 2011; CHEE *et al.*, 2011; FREIFELD *et al.*, 2014; GURULINGAPPA *et al.*, 2012; HUR *et al.*, 2012; MCCART *et al.*, 2013; ONG *et al.*, 2012; RAMESH *et al.*, 2014; ROCHEFORT *et al.*, 2015; ROSS *et al.*, 2013; YANG *et al.*, 2015).

A precisão de um classificador avalia o percentual de registros verdadeiramente positivos sobre o total de amostras classificadas como positivas. Já a sensibilidade de um classificador corresponde à taxa positiva verdadeira (TP) ou, em outras palavras, à proporção de registros positivos (possui ISC) corretamente atribuídos como positivos pelo classificador. Precisão e sensibilidade são definidas conforme Equações 6 e 7 (JAPKOWICZ, 2013):

(6)

$$\text{Precisão} = \frac{TP}{TP + FP}$$

(7)

$$\text{Sensibilidade} = \frac{TP}{TP + FN}$$

Onde *FN* é o número os registros falsos negativos, ou seja, quando uma ISC ocorreu e não foi constatada, e *FP* são falsos positivos, quando uma ISC não ocorreu e foi constatada pelo classificador. Já a ROC retém duas métricas durante a classificação, *TP* e *FPR* (taxa de falsos positivos), sendo que *FPR* é definida pela Equação 8 (JAPKOWICZ, 2013):

(8)

$$FPR = \frac{FP}{FP + TN}$$

Onde, *TN*, verdadeiros negativos, é o número de registros sem ISC classificados corretamente como negativos. Cada classificação *TP* e *FPR* é então inserida em um gráfico formando uma curva ROC para esse classificador, de modo que quanto mais próxima a curva está do canto superior esquerdo do espaço ROC (baixo *FP* e alto *TP*), melhor será o desempenho e a AUC desse classificador (JAPKOWICZ, 2013).

3.4.6 RESULTADOS

Dados de registros textuais, incluindo 27.648 descrições de cirurgias, foram analisados. Depois de excluir registros vazios e que não se adequassem aos critérios do estudo, conforme Tabela 2, o número de registros foi reduzido para 15.479, sendo 188 descrições cirúrgicas que tiveram ISC após a cirurgia e 15.291 descrições sem ISC constatada após a cirurgia. Conforme a Tabela 1, a proporção de registros analisados, 1,21% para 98,79%, reflete o cenário real das ISC do HCPA, sendo um conjunto de dados com classes desequilibradas. Do volume total de descrições cirúrgicas, 15.277 tiveram o potencial de contaminação informado, composto por 60,6% de cirurgias limpas, 32,63% cirurgias potencialmente contaminadas, 5,76% cirurgias contaminadas e 1,01% cirurgias infectadas, conforme Tabela 2.

Tabela 2. Registros textuais selecionados

Descrição	Cirurgias infectadas	Cirurgias não infectadas	Total
Número de documentos textuais disponibilizados	247	27401	27648
Registros sem texto	-29	-12103	-12132
Registros de pacientes que realizaram mais de uma cirurgia no mesmo período, das quais pelo menos uma não infectou	-4	-7	-11
Registros de infecções constatadas 30 (trinta) dias após a cirurgia	-26	0	-26
Número de registros utilizados no estudo	188 (1,21%)	15291 (98,79%)	15479 (100%)

As infecções do local cirúrgico são detectadas pela CCIH de maneira metódica, 3 vezes por semana. Relatórios de antibióticos, resultados de exames microbiológicos e demais prescrições são utilizados para realizar a triagem das evoluções que serão analisadas e discutidas para verificar a existência de ISC. A partir da constatação de infecção, é possível saber quais cirurgias vieram a ocasionar uma infecção, possibilitando a investigação do processo cirúrgico na busca por relações causais entre o processo e a infecção.

A Tabela 3 apresenta o desempenho de cada método de classificação. O melhor desempenho, considerando a relação entre verdadeiros positivos e falsos positivos representado pela área sob a curva ROC, foi obtido pelo método Stochastic Gradient Descent (SGD). Esse resultado foi atingido utilizando 75% dos termos selecionados a partir do teste X^2 , em um conjunto de termos TF sem normalização, com peso de 0,01 para a classe negativa.

Tabela 3. Desempenho dos classificadores. SA: Seleção de Atributos, P: Percentil, T: Transformação, N: Normalização, CW: Peso / Probabilidade Prévia, DP: Desvio Padrão.

Método	SA	P	T	N	CW	Precisão		Sensibilidade				ROC-AUC			
						Positivos	Negativos	Positivos	Negativos	Média	DP	Média	DP		
AdaBoost	X^2	20%	TF-IDF	-	-	2,2%	0,2%	99,7%	0,2%	89,2%	8,2%	51,%	1,5%	70,1%	4,1%
Decision Tree	X^2	25%	TF	-	1/0,01	3,1%	0,9%	99,4%	0,2%	60,6%	13%	75,5%	4,4%	68,1%	6,7%
Random Forest	F-escore	85%	TF-IDF	I1	1/0,01	2,8%	0,2%	99,8%	0,1%	89,9%	6,2%	62,7%	1,8%	76,3%	3,3%
Logistic Regression	F-escore	55%	TF-IDF	I1	1/0,005	2,6%	0,2%	99,9%	0,1%	95,7%	4,8%	56,2%	1,6%	75,9%	2,5%
LinearSVC	F-escore	85%	TF	I1	1/0,005	4,5%	0,6%	99,7%	0,1%	78,7%	9,6%	79,3%	1,3%	79%	4,7%
SVC	F-escore	10%	TF	I2	1/0,001	4,5%	99,5%	99,5%	0,1%	69,1%	9,9%	81,5%	3,3%	75,3%	4,6%
Nearest Centroid	X^2	20%	TF	I1	-	3,8%	0,4%	99,7%	0,1%	82,3%	8,9%	74,1%	1,4%	78,2%	4,3%
SGD	X^2	75%	TF	-	1/0,01	3,5%	0,3%	99,8%	0,1%	89,3%	6,8%	65%	1,4%	79,7%	3,3%
Multinomial Naive Bayes	X^2	40%	TF	-	20/80	6,4%	1,1%	99,5%	0,1%	61,1%	9,1%	88,9%	0,8%	75%	4,6%

A estratégia de classificação e pré-processamento do método SGD atingiu uma média de 79,7% de ROC-AUC (DP 3,3%), Sensibilidade média de 89,3% (DP 6,8%) para classes positivas e 65% (DP 1,4%) para classes negativas. Dado o

desequilíbrio entre as classes, os valores médios de precisão obtidos foram de 3,5% (DP 0,3%) para a classe positiva e 99,8% (DP 0,1%) para a classe negativa.

As estratégias de pré-processamento de cada método foram determinadas a partir de uma grade com diversas opções de parâmetros na qual os atributos SA, P, T, N e CW foram selecionados por terem obtido a melhor ROC-AUC, a qual é apresentada na Figura 5. A busca pelos melhores parâmetros foi realizada por meio de uma busca aleatória em função do seu desempenho e custo computacional (BERGSTRA e BENGIO, 2012).

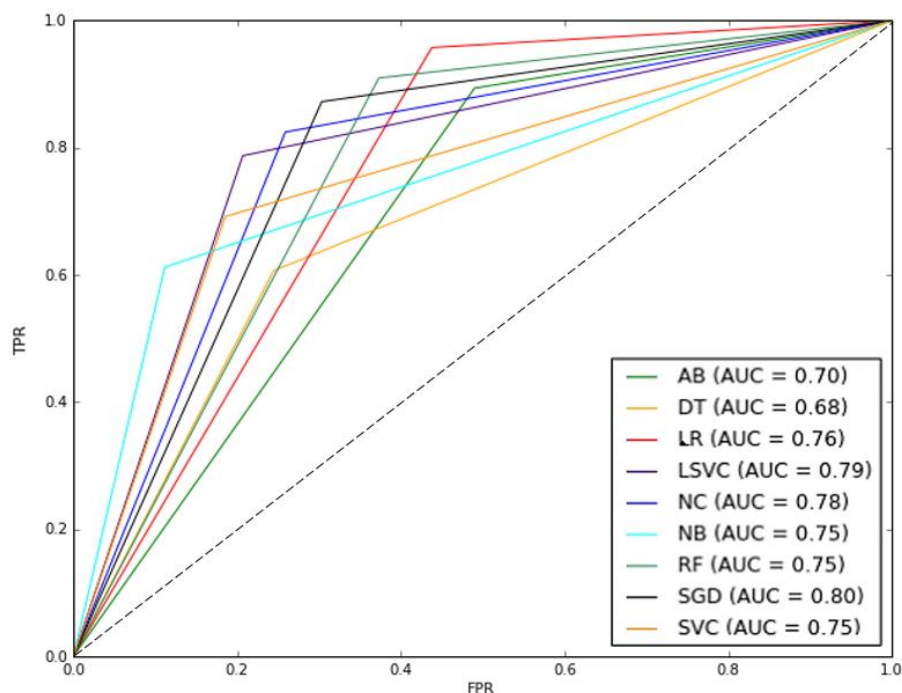


Figura 5. ROC-AUC: AD: AdaBoost, DT: Decision Tree, LSVC: LinearSVC, LR: *Logistic Regression*, NB: MultinomialNB, NC: *Nearest Centroid*, RF: *Random Forest*, SGD: *Stochastic Gradient Descent*, SVC: Support Vector Classification

A mineração de textos apresenta potencial para apoiar a previsão de ISC podendo direcionar o acompanhamento de pacientes com maior risco de infecção baseado na descrição cirúrgica, buscando minimizar o número de ISC. A Figura 6 apresenta a proporção de Falsos Negativos e a redução do volume de descrições a serem analisadas a partir da utilização de cada método.

A redução do número de cirurgias a serem acompanhadas apresenta um ganho para a vigilância de ISC em termos de redução de esforço e custo, porém o impacto das cirurgias que tiveram infecção e não foram constatadas (FN) pelos métodos deve ser considerado. Uma maior sensibilidade positiva pode ser obtida em

alguns métodos como no Logistic Regression, que atingiu uma média de 95,7% de sensibilidade (DP 4,8%), porém, dado o maior número de falsos positivos, o volume de cirurgias que irão requerer acompanhamento preventivo de ISC será maior.

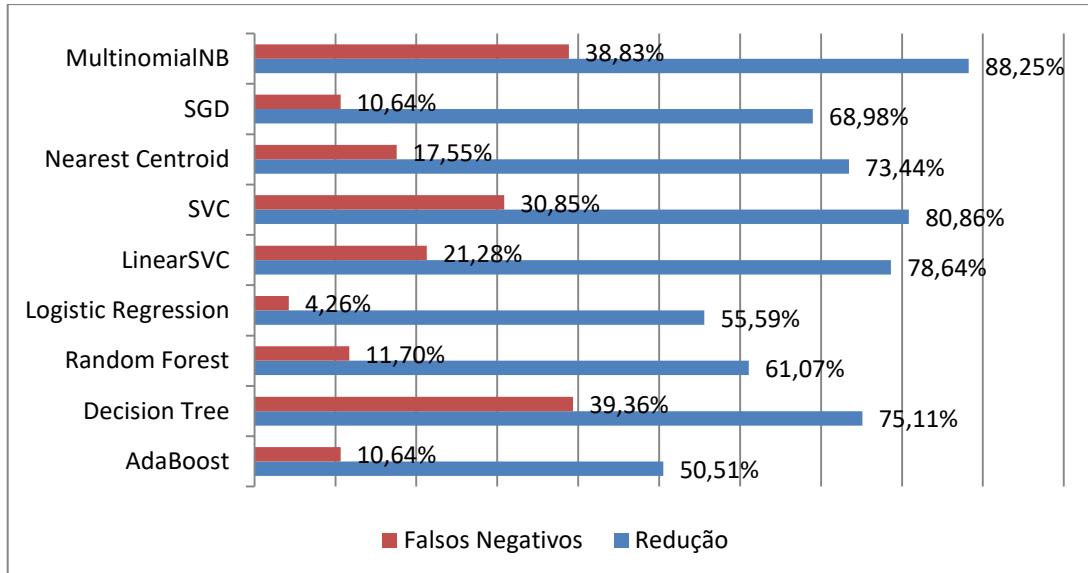


Figura 6. Redução do volume de dados vs. Falsos Negativos.

Das 188 descrições cirúrgicas positivas analisadas, 183 tinham o potencial de contaminação informado, sendo 79 cirurgias limpas (43%), 67 potencialmente contaminadas (37%) e 37 cirurgias contaminadas (20%). A partir do método SGD, 88,3% das cirurgias limpas, 86,4% das cirurgias potencialmente contaminadas e 100% das cirurgias contaminadas que tiveram ISC após a cirurgia foram previstas corretamente, totalizando 168 descrições cirúrgicas com ISC previstas, conforme Figura 7.

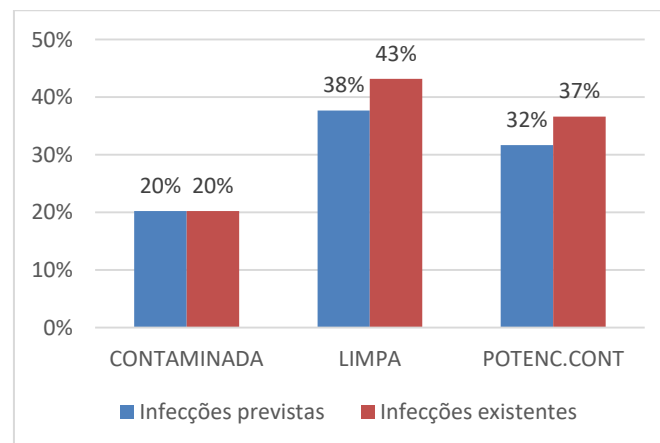


Figura 7. Cirurgias infectadas existentes e previstas.

Analisando também as 15.094 descrições cirúrgicas negativas com potencial de contaminação informado, o método SGD previu corretamente 70,9% das cirurgias limpas, 71,85% das cirurgias contaminadas, 50,98% das cirurgias infectadas e 46,5% das cirurgias contaminadas que não tiveram ISC. O baixo desempenho na previsão negativa das cirurgias infectadas e contaminadas deve-se a existência de termos comuns às cirurgias positivas presentes nos registros negativos. Os termos com maior valor preditivo para ISC foram, conforme esperado, derivados de “côlon”, “infecção”, entre outros, indicados durante a cirurgia.

O método SGD com as opções de pré-processamento descritas anteriormente foi escolhido o melhor método para previsão de infecções no texto da descrição de cirurgias (baseado na ROC-AUC). Porém, ao buscar a redução do volume de acompanhamento preventivo de ISC minimizando o número de falsos negativos (infecções não previstas pelo classificador), o método Logistic Regression deve ser considerado. Outras combinações de métodos e pré-processamento para previsão resultaram em escores ROC-AUC que não foram estatisticamente diferentes desses dois métodos, dado o alto desvio padrão em função do desequilíbrio das classes.

3.5 CONSIDERAÇÕES FINAIS

A MT pode ser usada para apoiar a previsão de infecções em pacientes cirúrgicos e possivelmente reduzir o número de ISC no período pós-operatório. A MT pode ser aplicada logo após a realização da cirurgia para apoiar a tomada de decisões de prevenção, como a prescrição de medicamentos, agendamento de consultas e também a identificação de fatores causadores de ISC.

A partir dos métodos de MT utilizados, é possível direcionar os esforços reduzindo 55,59% do volume de cirurgias a serem acompanhadas preventivamente com apenas 4,26% de infecções não previstas, ou, reduzir o volume de cirurgias a serem acompanhadas em 68,98% aceitando que 10,64% das infecções possam não ser previstas e evitadas. As 188 descrições cirúrgicas que tiveram ISC no período pós-operatório são baseadas na vigilância realizada pela CCIH, feita a partir de critérios que apontam a ocorrência de infecção. Esses critérios direcionam o trabalho de vigilância, mas podem não detectar todas as infecções existentes. Dessa forma,

podem haver cirurgias que verdadeiramente vieram a ter infecções previstas, mas foram classificadas como falso positivo.

Boa parte das descrições técnicas cirúrgicas positivas para ISC pós-operatória possuem texto pouco informativo e por vezes repetido em cirurgias negativas para ISC pós-operatória. Técnicas de pré-processamento, como a seleção de atributos e transformação TF-IDF, ajudam a superar esses desafios, porém, em larga escala, essa característica pode afetar negativamente a precisão e sensibilidade da classificação.

O presente estudo aponta métodos eficazes de MT para apoiar a previsão de ISC no período pós-operatório e pode ser aplicado em outros tipos de EA, proporcionando a melhoria dos cuidados e segurança do paciente. Esta pesquisa buscou, de maneira exploratória, avaliar o desempenho de métodos de pré-processamento e MT para identificar cirurgias que terão ISC no período pós-operatório.

Como trabalhos futuros, sugere-se a aplicação dos métodos de MT apenas em determinados grupos de pacientes, selecionados a partir de métodos de agrupamento ou em condições médicas específicas bem como a utilização de outras informações disponíveis no sistema informatizado do hospital. Sugere-se também a utilização de rótulos de classes gramaticais nas palavras, a análise de sentenças.

3.6 REFERÊNCIAS

AGGARWAL, C. C., ZHAI, C. Mining Text Data. Springer Science Business Media, 2012.

ANVISA. Critérios Diagnósticos de Infecções Relacionadas à Assistência à Saúde. Segurança do Paciente e Qualidade em Serviços de Saúde. Brasília, 2013.

ANVISA. Disponível em <http://www.anvisa.gov.br/servicosaude/controle/roteiro.htm#2>. Acessado em: 27/01/2017.

ARAMAKI, E. MIURA, Y. TONOIKE, M. OHKUMA, T. MASUICHI, H. WAKI, K. OHE, K. Extraction of adverse drug effects from clinical records. Studies in Health Technology and Informatics, 2010.

BERGSTRA, J., BENGIO, Y. Random search for hyper-parameter optimization. Journal of Machine Learning Research, 2012.

BROWN, K. L., RIDOUT, D. A., SHAW, M., DODKINS, I., SMITH, L. C., O'CALLAGHAN, M. A., ... & HARTLEY, J. C. Healthcare-associated infection in pediatric patients on extracorporeal life support: The role of multidisciplinary surveillance. Pediatric Critical Care Medicine, 2006.

CLASSEN, D. C., RESAR, R., GRIFFIN, F., FEDERICO, F., FRANKEL, T., KIMMEL, N., WHITTINGTON, J. C., FRANKEL, A., SEGER, A., JAMES, B. C. 'Global trigger tool' shows that adverse events in hospitals may be ten times greater than previously measured. Health affairs, 2011.

CURRAN, E. T., COIA, J. E., GILMOUR, H., MCNAMEE, S., & HOOD, J. Multi-centre research surveillance project to reduce infections/phlebitis associated with peripheral vascular catheters. Journal of Hospital Infection, 2000.

DALTOÉ, T., BREIER, A., SANTOS, H. B., WAGNER, M. B., KUCHENBECKER, R. S. Serviços de Controle de Infecção Hospitalar: características, dimensionamento e atividades realizadas. Revista da Sociedade Brasileira de Clínica Médica, 2014.

DEMNER-FUSHMAN, D. CHAPMAN, W.W. MCDONALD, C.J. What can natural language processing do for clinical decision support? Journal of Biomedical Informatics, 2009.

DONG, Y. X., LI, Y. X., LI, J. H., ZHAO, H. Analysis on weighted AUC for imbalanced data learning through isometrics. Journal of Computational Information, 2012.

EDWARDS, I.R.; BIRIELL, C. Harmonisation in pharmacovigilance. Drug Safety, 1994.

FAYYAD, U.; PIATETSKY-SHAPIRO, G; SMYTH, P. From data mining to knowledge discovery: An overview. In Advances in Knowledge Discovery and Data Mining, 1996.

FREEMAN, R., MOORE, L. S., ÁLVAREZ, L. G., CHARLETT, A., & HOLMES, A. Advances in electronic surveillance for healthcare-associated infections in the 21st Century: a systematic review. *Journal of Hospital Infection*, 2013.

GARNER, J. S. JARVIS, W.R. EMORI, T.G. HORAN, T.C. HUGHES, J.M. CDC definitions for nosocomial infections. *American Journal of Infection Control*, 1988.

GRIFFIN, F.A. RESAR, R.K. IHI Global Trigger Tool for measuring adverse events (2nd edition). IHI Innovation Series white paper, 2009.

HALEY, R. W., CULVER, D. H., WHITE, J. W., MORGAN, W. M., EMORI, T. G., MUNN, V. P., & HOOTON, T. M. The efficacy of infection surveillance and control programs in preventing nosocomial infections in us hospitals. *American journal of epidemiology*, 1985.

HAN, J. PEI, J., KAMBER, M. Data Mining: concepts and techniques (2nd edition). Morgan Kaufmann, 2006.

HCPA. Disponível em: <https://www.hcpa.edu.br/content/view/7925/2429/>. Acessado em: 18/02/2017.

KANG, N. AFZAL, Z. SINGH, B. MULLIGEN, E. M. V. KORS, J. A. Using an ensemble system to improve concept extraction from clinical records. *Journal of Biomedical Informatics*, 2012.

MANGRAM, A. J., HORA, T. C., PEARSON, M. L., SILVER, L. C., JARVIS, W. R. Guideline for prevention of surgical infection, 1999. *Infect Infection Control & Hospital Epidemiology*, 1999.

MCCART, J. A. BERNDT, D. J. JARMAN, J. FINCH, D. K. LUTHER, S. L. Finding falls in ambulatory care clinical documents using statistical text mining. *Journal of American Medical Informatics Association*, 2013.

MELLING, A. C., ALI, B., SCOTT, E. M., & LEAPER, D. J. Effects of preoperative warming on the incidence of wound infection after clean surgery: a randomised controlled trial. *The Lancet*, 2001.

MEYSTRE, S.M. SAVOVA, G.K. KIPPER-SCHULER, K.C. HURDLE, J.F. Extracting information from textual documents in the electronic health record: a review of recent research. *Yearbook Medical Informatics*, 2008.

MICHELSON, J. Improved detection of orthopaedic surgical site infections occurring in outpatients. *Clinical orthopaedics and related research*, 2005.

MICHELSON, J. D.; PARISEAU, J. S.; PAGANELLI, W. C. Assessing surgical site infection risk factors using electronic medical records and text mining. *American journal of infection control*, 2014.

RANJI, S. R., SHETTY, K., POSLEY, K. A., LEWIS, R., SUNDARAM, V., GALVIN, C. M., & WINSTON, L. G. Closing the Quality Gap: A Critical Analysis of Quality Improvement Strategies. Vol. 6: Prevention of Healthcare-Associated Infections, 2007.

ROCHA, J. J. R. Infecção em cirurgia e cirurgia das infecções. Medicina - Ribeirão Preto, 2008.

SANTOS, A. C. R. B. D. MOTA, D. M. GOMES, S. M. Investigação de eventos adversos em serviços de saúde. (1ª edição). Agência Nacional de Vigilância Sanitária, 2013.

SILVA, A. CORTEZ, P. SANTOS, M. F. GOMES, L. NEVES, J. Rating organ failure via adverse events using data mining in the intensive care unit. Artificial Intelligence in Medicine, 2008.

STONE, Patricia W.; KUNCHES, Lauren; HIRSCHHORN, Lisa. Cost of hospital-associated infections in Massachusetts. American journal of infection control, 2009.

TAN, A. Text Mining: The state of the art and the challenges. Proceedings of the PAKDD 1999 Workshop on Knowledge Discovery from Advanced Databases, 1999.

TOWNSEND, C., BEAUCHAMP, R. D., EVERS, B. M., & MATTOX, K. L. SABISTON. Tratado de Cirurgia: a base biológica da prática cirúrgica moderna. Elsevier Brasil, 2014.

WATCHER, R. M. Compreendendo a Segurança do Paciente – 2. ed. Mcgraw-Hill Education, 2013.

YOKOE, D. S., MERMEL, L. A., ANDERSON, D. J., ARIAS, K. M., BURSTIN, H., CALFEE, D. P., COFFIN, S. E., DUBBERKE, E. R., FRASER, V., GERDING, D. N., *et al.* Executive summary: a compendium of strategies to prevent healthcare-associated infections in acute care hospitals. Infection Control & Hospital Epidemiology, 2008.

ZAFARANI, R., ABBASI, M. A., LIU, H. Social Media Mining, An Introduction. Cambridge University Press, 2014.

4 ARTIGO 3: Detecção de Infecções do Sítio Cirúrgico por meio da mineração de textos no Hospital de Clínicas de Porto Alegre

Resumo: O presente estudo foi realizado no Hospital de Clínicas de Porto Alegre (HCPA), Brasil, que utiliza sistemas informatizados e prontuários eletrônicos, buscando verificar a hipótese de que o uso de métodos de mineração de texto pode ser uma alternativa viável para detecção de Infecções em Sítio Cirúrgico (ISC). Identificar e medir a ocorrência de ISC é fundamental para a melhoria da segurança em hospitais. O conhecimento exato das taxas de ISC a partir da sua detecção pode ser usado como parte do mecanismo de *feedback* para diminuir a incidência futura de infecções. Nesse contexto, este estudo teve por objetivo avaliar o desempenho de métodos de mineração de textos na detecção de ISC no período pós-operatório. Foram selecionados 12.637 prontuários eletrônicos, nos quais foram aplicadas técnicas de pré-processamento de texto e os métodos de classificação LinearSVC, Logistic Regression, Multinomial Naive Bayes, Nearest Centroid, Random Forest, Stochastic Gradient Descent, Support Vector Classification (SVC), possibilitando a comparação com os resultados da vigilância já realizada no HCPA. O algoritmo Logistic Regression, apresentou o melhor desempenho 80,6% pela métrica ROC-AUC (Receiver Operating Characteristic – Area Under the Curve), já o pior desempenho foi obtido pelo método SVC, 61%. Os métodos de mineração de textos podem ser usados para apoiar diariamente a detecção de ISC, bem como orientar melhorias na segurança do paciente no período pós-operatório.

Palavras-chave: Detecção de Eventos Adversos; Infecção do Sítio Cirúrgico; Mineração de Texto.

4.1 INTRODUÇÃO

A ocorrência de Eventos Adversos (EA) é considerada um problema internacional e sua detecção é uma tarefa complexa. EA geralmente são constatados por meio da notificação voluntária dos profissionais da saúde ou por meio da revisão e análise de prontuários por amostragem (GRIFFIN e RESAR, 2009). Esses EA podem estar relacionados a diversas causas, boa parte já documentada na literatura,

porém a sua constatação demanda grande esforço por parte dos profissionais da saúde.

A identificação das causas e a prevenção de EA podem trazer benefícios em termos de custos, diminuição da mortalidade e também diminuição do tempo de tratamento do paciente (SANTOS *et al.*, 2013). A utilização de sistemas de informação para identificação de EA vem sendo citada na literatura como um assunto emergente, principalmente relacionada à análise de textos clínicos (KANG *et al.*, 2012; MEYSTRE *et al.*, 2008; DEMNER-FUSHMAN *et al.*, 2009).

No contexto cirúrgico, as Infecções em Sítio Cirúrgico (ISC) são um dos principais EA a serem evitados e a Comissão de Controle de Infecção Hospitalar (CCIH) é o órgão encarregado de informar, vigiar, pesquisar, prevenir e controlar as infecções hospitalares (ANVISA, 2017). O objetivo principal de um programa de controle de infecção é exercer uma ação educativa na comunidade hospitalar quanto à importância da prevenção e do controle das infecções e, assim, reduzir o nível de infecção hospitalar endêmica e epidêmica.

Nas instituições de saúde, as informações do paciente são armazenadas principalmente em textos narrativos e relatórios clínicos. Dada a crescente utilização de registros eletrônicos hospitalares, é comum que um paciente possua centenas de notas clínicas registradas por diferentes áreas do hospital (FRIEDMAN e ELHADAD, 2014). Nesse sentido, técnicas como o Processamento de Linguagem Natural (PLN), Mineração de Textos (MT), Aprendizagem de Máquina (AM), entre outras, buscam facilitar de maneira automatizada a representação das informações relevantes no texto.

Os avanços das tecnologias em hardware e software têm possibilitado a rápida criação de repositórios de diferentes tipos de dados. Com o aumento da quantidade de texto disponível em diferentes aplicações, tem crescido a necessidade de métodos que possam aprender padrões interessantes dos dados. (AGGARWAL e ZHAI, 2012). Nesse contexto, o presente estudo tem por hipótese que a MT pode ser usada para detecção de infecções decorrentes da cirurgia a partir de texto livre de registros médicos do período pós-operatório, sendo o objetivo do presente estudo avaliar o desempenho de métodos de mineração de textos na detecção de infecções no período pós-operatório.

O presente artigo está organizado em cinco seções além dessa introdução. A seção 4.2 aborda o referencial teórico da pesquisa e as seções 4.3 e 4.4 apresentam

os procedimentos metodológicos e o processo de MT. A seção 4.5 apresenta as considerações finais.

4.2 REFERENCIAL TEÓRICO

Nos últimos anos, tem crescido a utilização de ferramentas de rastreamento como método para medir a ocorrência de EA na saúde (WATCHER, 2013). As sensibilidades de métodos para detecção de casos de infecção no sítio cirúrgico a partir de registros eletrônicos relatadas na literatura variam de 60% a 98% (FREEMAN *et al.*, 2013), sendo que os métodos com maior sensibilidade são aplicados em áreas e conjuntos de dados específicos da saúde.

Um dos princípios-chave de segurança do paciente está centrado em aprender com os próprios erros, sendo que a discussão aberta desses erros faz parte da cultura de sistemas seguros. A utilização de métodos de rastreamento e detecção tem por hipótese que erros no cuidado da saúde geram indícios, ou seja, EA, os quais podem ser encontrados para posterior análise de causa raiz. A ferramenta mais popular é a Ferramenta Global de Rastreamento (FGR), desenvolvida pelo Institute for Healthcare Improvement (IHI) (WATCHER, 2013). FGR trata-se de um conjunto de gatilhos que indicam a ocorrência de EA, servindo para guiar a revisão manual de prontuários e demais dados clínicos do paciente.

Apesar de as ferramentas de rastreamento não terem uma sensibilidade e especificidade ótimas na identificação de EA, elas apresentam uma boa confiabilidade (SHAREK *et al.*, 2011). Muitas das situações detectadas não apresentam risco ou dano real, mas podem ser utilizadas como triagem para uma revisão mais detalhada do prontuário (WATCHER, 2013), reduzindo o número de prontuários a serem analisados, bem como o tempo e o custo das ações de melhoria da segurança do paciente.

Entre as recentes abordagens para detecção de EA, estão as que utilizam técnicas de mineração de dados (MD), a qual é parte do processo chamado de Descoberta de Conhecimento em Bases de Dados (DCBD). DCBD se refere a todo o processo de descoberta de conhecimento útil de bases de dados, sendo descrito como o processo não trivial de identificação de padrões válidos, úteis, e compreensíveis nos dados (FAYYAD, 1996). A MD, que tem como base algoritmos

de AM, é uma importante etapa do processo de DCBD e que muitas vezes é utilizada por si só para descrever a extração de conhecimento de bases de dados (HAN *et al.*, 2006).

Algoritmos de mineração de dados podem ser divididos em duas categorias: aprendizagem supervisionada e aprendizagem não supervisionada. Na aprendizagem supervisionada os registros analisados possuem classes conhecidas e é tarefa do algoritmo identificar características que possibilitem diferenciar os registros entre as classes. A aprendizagem supervisionada pode ser dividida em classificação e regressão (ZAFARANI *et al.*, 2014).

De maneira geral, qualquer método que incorpore informações de amostras de treinamento na concepção de um classificador emprega o aprendizado. Nesse sentido, aprendizagem refere-se a alguma forma de algoritmo para reduzir o erro de um modelo em um conjunto de dados de treinamento (DUDA *et al.*, 2012). Para apoiar a detecção de eventos adversos, estudos têm utilizado algoritmos de AM na mineração de dados em várias áreas da saúde, descobrindo padrões de sequências, associações e agrupamentos em prontuários eletrônicos, apoiando também a previsão de eventos adversos e identificação das áreas associadas a eles (ARAMAKI *et al.*, 2010; MCCART *et al.*, 2013; SILVA *et al.*, 2008).

Já a MT refere-se ao processo de extração de padrões de interesse do usuário, não triviais, a partir de documentos de texto não estruturado (TAN, 1999). A MT pode utilizar algoritmos de AM, que podem explorar dados de treinamento para aprender a classificar ou prever classes não vistas antes (AGGARWAL e ZHAI, 2012). Identificação de eventos adversos a partir de textos clínicos pode ser considerada um problema de atribuição dos registros textuais à classe correta, por exemplo, possui ou não possui infecção. Dessa forma pode-se avaliar estatisticamente a pureza das classificações por meio dos verdadeiros e falsos positivos e negativos. Porém, conforme Penz *et al.* (2007) afirmam, deve-se considerar que erros ortográficos e abreviaturas são uma clara fonte de erro e diminuem a precisão dos métodos.

No contexto cirúrgico, MT pode ser usada no texto clínico do paciente inserido após a cirurgia, bem como no período de internação ou nas consultas de retorno do paciente durante os 30 (trinta) dias após a cirurgia. Esse texto pode conter pistas importantes que revelem a ocorrência de infecções, como a utilização de antibióticos, febre ou o próprio relato de ISC. Esse texto será transformado em dados estruturados, nos quais algoritmos e métodos de análises quantitativas podem ser aplicados para

detectar padrões relacionados à ocorrência de infecções e então classificar novas ocorrências. O processo de MT pode pré-visualizar notas para reduzir o volume de textos revistos manualmente, direcionando o trabalho de vigilância conduzido pela CCIH, por exemplo.

A ISC após a cirurgia é determinada por uma combinação de fatores: quantidade e tipo de contaminação, técnica cirúrgica e anestésica empregada e resistência do hospedeiro. Essas e outras definições da literatura compõem o que caracterizam as ISC no Hospital de Clínicas de Porto Alegre (HCPA) (CULVER *et al.*, 1991; GARNER *et al.*, 1988; GAYNES *et al.*, 2001; MANGRAM, *et al.*, 1999), que são determinadas, também, pelos seguintes indícios, geralmente, encontrados no texto clínico: (i) drenagem purulenta da incisão superficial; (ii) cultura positiva de fluidos ou tecidos da incisão superficial, obtida com técnica asséptica; (iii) dor ou sensibilidade, inflamação local, vermelhidão ou calor; (iv) diagnóstico médico de infecção cirúrgica superficial.

Também é possível perceber na literatura que a utilização de técnicas de MT pode e vem trazendo benefícios para apoiar a análise de infecções (BOUZBID *et al.*, 2011; MICHELSON *et al.*, 2014; CAMPILLO-GIMENEZ *et al.*, 2013). No entanto se faz necessário o desenvolvimento de pesquisas que verifiquem a aplicabilidade e proponham melhorias aos métodos para detecção de EA e apoiem o controle de infecções hospitalares no contexto dos hospitais brasileiros (DALTOE *et al.*, 2014).

4.3 PROCEDIMENTOS METODOLÓGICOS

Este estudo tem natureza aplicada e busca gerar soluções para a detecção de infecções no período pós-operatório. Este estudo também tem caráter descritivo, pois visa apresentar como métodos de MT podem ser aplicados para apoiar a detecção de ISC, utilizando abordagens quantitativas para análise e definição de soluções.

Foram aplicados métodos de MT para avaliar registros clínicos de um hospital brasileiro com a intenção de detectar ISC no período pós-operatório. Foram utilizados dados disponibilizados pelo hospital para comparar técnicas de pré-processamento e classificação de texto. Trata-se de um estudo retrospectivo que utiliza métodos de MT para detectar ISC com base em registros médicos gerados em até 30 (trinta) dias após

uma cirurgia. O estudo foi realizado em quatro etapas que contemplam desde a seleção dos dados do sistema informatizado do hospital até a avaliação de desempenho dos classificadores, conforme Figura 7.

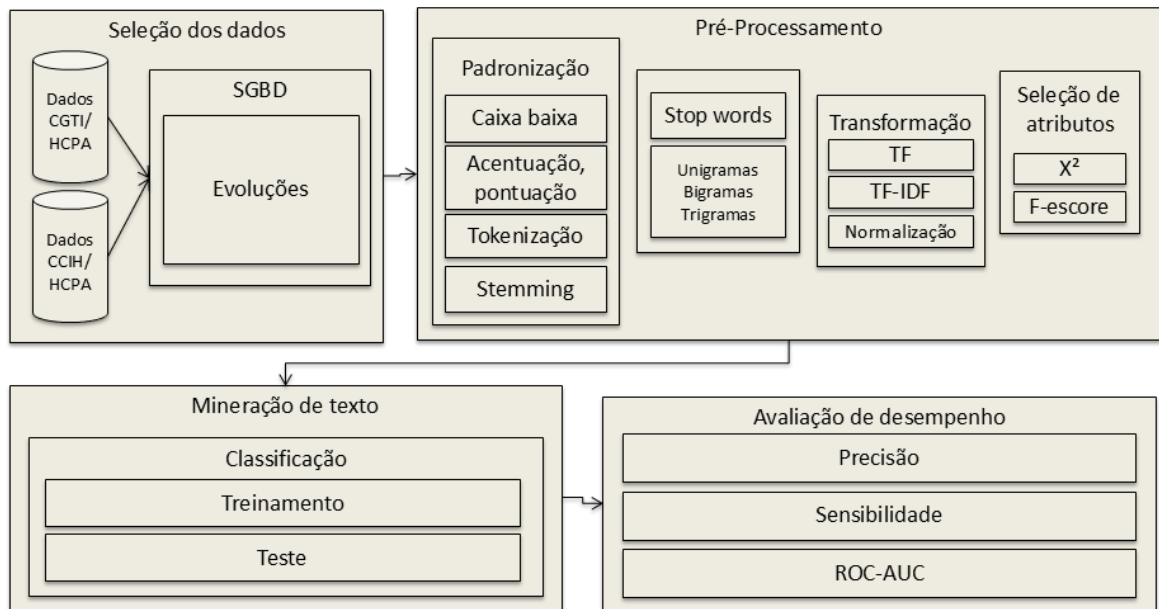


Figura 8. Procedimentos Metodológicos

Para gerenciar os dados do estudo foi utilizado o Sistema Gerenciador de Banco de Dados (SGBD) PostgreSQL, versão 9.6 (POSTGRESQL, 2017). Já para o pré-processamento, aplicação de métodos de MT e avaliação de desempenho foi utilizada a linguagem de programação de alto nível, Python versão 3.5 (PYTHON, 2017). O Python inclui bibliotecas para apoiar o processamento dos dados, como Natural Language Toolkit (NLTK) (BIRD, *et al.* 2009) e Scikit-learn (PEDREGOSA *et al.*, 2011), que foram utilizadas neste estudo.

A etapa de seleção dos dados teve por objetivo recuperar e integrar os dados da CCIH e das evoluções dos pacientes contidas nos prontuários eletrônicos do sistema informatizado do HCPA, preparando-os para o pré-processamento. A Comissão de Controle de Infecções já identifica infecções por meio da revisão retrospectiva dos prontuários. Os dados históricos já analisados no HCPA possibilitaram o treinamento de algoritmos de classificação bem como a avaliação do desempenho da utilização de técnicas de MT frente ao trabalho já realizado no hospital.

O pré-processamento buscou preparar e estruturar o texto da descrição cirúrgica para as análises quantitativas dos métodos de classificação. Nesta etapa foi

realizada a padronização do texto, que inclui transformação do texto para caixa baixa, remoção de acentuação e pontuação, tokenização, *stemming*, remoção de *stop words*, e criação de unigramas, bigramas e trigramas, bem como os processos de transformação e seleção de atributos, que serão descritos detalhadamente nas seções a seguir.

Já na etapa da mineração de texto propriamente dita foram aplicados os métodos de classificação Linear Support Vector Classification (LinearSVC), Logistic Regression, MultinomialNB, Nearest Centroid, Random Forest, Stochastic Gradient Descent (SGD), Support Vector Classification (SVC), que a partir do treinamento e teste com validação cruzada, identificaram prontuários como positivos ou negativos para ISC. A seleção de parâmetros e estratégia de pré-processamento ocorreu por meio da busca aleatória em uma grade de parâmetros buscando encontrar a melhor combinação de parâmetros para cada método (BERGSTRA e BENGIO, 2012).

Os métodos de classificação foram selecionados a partir de sua disponibilidade na biblioteca Scikit-learn (PEDREGOSA *et al.*, 2011) e utilização em estudos anteriores: (i) Support Vector Machines (SVM) (ARAMAKI *et al.*, 2010; BIAN *et al.*, 2012; BOTSIS *et al.*, 2011; BRUIJN *et al.*, 2011; CHEE *et al.*, 2011; HUR *et al.*, 2012; MCCART *et al.*, 2013; GENKIN *et al.*, 2007; ONG *et al.*, 2012; RAMESH *et al.*, 2014; ROCHEFORT *et al.*, 2015; ROSS *et al.*, 2013; YANG *et al.*, 2015); (ii) Logistic Regression (MCCART *et al.*, 2013; GENKIN *et al.*, 2007); (iii) Naive Bayes (BOTSIS *et al.*, 2011; GURULINGAPPA *et al.*, 2012; ONG *et al.*, 2012; RAMESH *et al.*, 2014; ROSS *et al.*, 2013); (iv) Decision Trees (BOTSIS *et al.*, 2011; FREIFELD *et al.*, 2014; ROSS *et al.*, 2013); (v) Random Forest, Nearest Neighbors (BOTSIS *et al.*, 2011).

Por fim, a avaliação do desempenho dos classificadores foi feita com base na sua Precisão, Sensibilidade e a Área Sob a Curva das Características Operacionais do Receptor (ROC-AUC). Foram mantidos na análise apenas os métodos que atingiram mais de 60% de ROC-AUC após a otimização de parâmetros.

4.4 MINERAÇÃO DE TEXTOS

4.4.1 SELEÇÃO DOS DADOS

O HCPA utiliza o sistema informatizado AGH (Aplicativos de Gestão Hospitalar) no qual todo o texto clínico inserido por médicos, enfermeiros e acadêmicos fica centralizado, viabilizando a coleta e posterior aplicação de métodos de MT. O texto clínico dos pacientes foi disponibilizado pelo HCPA em planilhas do Microsoft Office Excel (.xlsx) onde cada linha corresponde ao texto das evoluções de um paciente 30 (trinta) dias após uma cirurgia. Esses registros foram então inseridos no Postgres, onde foi criada uma tabela para inserção dos dados consolidados, excluindo registros vazios, duplicados ou com inconsistências.

A tabela analisada foi formada por duas colunas, uma contendo o texto com a descrição da cirurgia e outra com um rótulo binário - sendo 1 para ocorrência de ISC e 0 para a não ocorrência de ISC - utilizado para classificar os registros. O rótulo dos registros infectados foi fornecido pela CCIH, que acompanha mensalmente o número de infecções pós-operatórias no HCPA.

O texto coletado de cada cirurgia diz respeito a todas as narrativas clínicas inseridas até 30 (trinta) dias depois da cirurgia, incluindo notas clínicas no período de internação ou retorno do paciente para uma consulta. A identificação de ISC refere-se a análise de um conjunto de dados com classes desbalanceadas, onde, inicialmente, 98,4% dos registros possuem rótulo 0 (negativo para ocorrência de ISC) e apenas 1,6% com o rótulo 1 (positivo para ocorrência de ISC).

Foram excluídos da análise registros em que o registro do paciente não possui texto no período analisado, como quando o paciente não é internado e não retorna para uma consulta durante 30 (trinta) dias após a cirurgia. Da mesma forma, foram excluídos os registros em que a ISC foi constatada com mais de 30 (trinta) dias após a cirurgia, pois o texto presente no período poderia não estar indicando a ocorrência de infecção. Também foram excluídos registros de pacientes que realizaram mais de uma cirurgia em menos de 30 (trinta) dias, das quais ao menos uma foi infectada e uma não foi infectada, caso esse em que não seria possível diferenciar a priori o texto relacionado ou não à ocorrência de ISC.

4.4.2 PRÉ-PROCESSAMENTO

O pré-processamento teve por objetivo preparar o texto para a aplicação dos métodos de classificação. Apenas dados do campo textual foram tratados nessa etapa. Assim como realizado no estudo anterior (SILVA *et al.*, 2017) a primeira etapa do pré-processamento buscou remover a acentuação das palavras, substituir letras maiúsculas por minúsculas e realizar a tokenização do texto (FELDMANN e SANGER, 2006). Após a tokenização, utilizando a biblioteca NLTK em português, foi possível realizar a remoção das *stop words* e também realizar o processo de normalização morfológica chamada de *stemming* utilizando o algoritmo RSLPStemmer (ORENGO e HUYCK, 2001), disponível no NLTK.

Após as etapas de padronização, tokenização, remoção de *stop words* e *stemming*, cada registro se tornou uma lista de *tokens* seguidos pelo rótulo da sua classe, 0 ou 1. A partir disso, foi realizado o processo de definição de unigramas, bigramas ou trigramas (PERKINS, 2014), pois cada um desses veio a ser um atributo analisado. Na etapa seguinte foi realizada a conversão do texto em uma matriz esparsa de contagem de termos (*tokens*), que possibilita a criação de análises quantitativas (SCIPY, 2017).

Nesse estudo foi utilizada a abordagem do saco de palavras, na qual a ordem exata dos termos em um documento é ignorada, mas o número de ocorrências de cada termo é mantido (MANNING, *et al.*, 2009). A Frequência do Termo (TF) e também a frequência ponderada pela frequência inversa do termo nos documentos (TF-IDF) foram analisadas (SKLEARN, 2017a). Opções de normalização da TF e TF-IDF foram analisadas a partir das normas ℓ_1 e ℓ_2 (YI e FALOUTSOS, 2000), buscando identificar a melhor opção de transformação para cada método de MT.

Após a transformação e normalização dos termos foi feita a seleção de atributos com base nas métricas qui-quadrado (X^2) e F-escore pela sua adequação ao problema e utilização na seleção de atributos em estudos anteriores (MICHELSON *et al.*, 2014; LUCINI, *et al.*, 2017), onde valores superiores indicam que os termos são mais determinantes para a classificação (SEBASTIANI 2002; GUYON *et al.*, 2006). Foram selecionados os atributos com base em um percentil dos maiores valores, sendo que o percentil e a métrica utilizada variaram para cada método de MT analisado.

4.4.3 CLASSIFICAÇÃO

Neste trabalho foram analisados vários métodos de classificação supervisionada, nos quais os valores de atributo da classe para o conjunto de dados são conhecidos antes de executar o algoritmo. Na aprendizagem supervisionada, os registros do conjunto de dados são tuplas no formato (x, y) , onde x é um vetor de termos tratados no pré-processamento e y é o atributo de classe, 0 ou 1 para a ocorrência de infecção, sendo tarefa do classificador encontrar um mapeamento $m()$ tal que $m(x) = y$ (ZAFARANI *et al.*, 2014).

O presente estudo apresenta um desequilíbrio na distribuição das classes alvo: 98,4% para amostras negativas e 1,6% para amostras positivas para ocorrência de ISC. Assim como no estudo anterior (SILVA *et al.*, 2017), foi utilizado o método de validação cruzada estratificada (10 partições) para garantir que o modelo aprendido por cada algoritmo retivesse maior aproveitamento sobre a classe minoritária (positiva) (SKLEARN, 2017b). Na validação cruzada estratificada, as classes são representadas em cada partição nas mesmas proporções que no conjunto de dados completo (98,4% e 1,6%). Por fim calcula-se a média dos 10 escores para obter uma estimativa global (WITTEN *et al.*, 2011).

Parâmetros que não são diretamente aprendidos pelo classificador, por exemplo C, kernel e gama para SVM (SKLEARN, 2017c), bem como o percentil, métrica (χ^2 e F-escore) para seleção de atributos, utilização de TF, TF-IDF, norma l_1 ou l_2 e também o peso das classes foram definidos por meio da busca aleatória em uma grade de parâmetros (BERGSTRA e BENGIO, 2012). Essa busca teve por objetivo encontrar a melhor combinação de parâmetros para maximizar a Área Sob a Curva das Características Operacionais do Receptor (ROC-AUC) (DONG *et al.*, 2012).

4.4.4 AVALIAÇÃO DO DESEMPENHO

A avaliação do desempenho dos classificadores foi feita a partir das métricas Precisão, Sensibilidade (*Recall*) e ROC-AUC. Adicionalmente matrizes de confusão foram analisadas. Essas métricas foram selecionadas pela sua adequação ao problema de classificação em conjuntos de dados desbalanceados e pela sua utilização em estudos anteriores (ARAMAKI *et al.*, 2010; BIAN *et al.*, 2012; BOTSIS *et*

al., 2011; BRUIJN *et al.*, 2011; CHEE *et al.*, 2011; FREIFELD *et al.*, 2014; GURULINGAPPA *et al.*, 2012; HUR *et al.*, 2012; MCCART *et al.*, 2013; ONG *et al.*, 2012; RAMESH *et al.*, 2014; ROCHEFORT *et al.*, 2015; ROSS *et al.*, 2013; YANG *et al.*, 2015).

A precisão de um classificador avalia a proporção de exemplos verdadeiramente positivos sobre o total de amostras classificadas como positivas. Já a Sensibilidade de um classificador corresponde à sua taxa positiva verdadeira (TP) ou, em outras palavras, à proporção de registros positivos (possui ISC) corretamente atribuídos como positivos pelo classificador. Precisão e Sensibilidade são apresentadas nas Equações 1 e 2.

(1)

$$\text{Precisão} = \frac{TP}{TP + FP}$$

(2)

$$\text{Sensibilidade} = \frac{TP}{TP + FN}$$

Já a curva ROC retém duas métricas durante a classificação, TP e FPR (taxa de falsos positivos), conforme definido na Equação 3 (JAPKOWICZ, 2013):

(3)

$$FPR = \frac{FP}{FP + TN}$$

Cada classificação TP e FPR é então inserida em um gráfico formando uma curva ROC para esse classificador, de modo que quanto mais próxima a curva está do canto superior esquerdo do espaço ROC (baixo *FP* e alto *TP*), melhor será o desempenho e maior será a AUC desse classificador (JAPKOWICZ, 2013).

4.4.5 RESULTADOS

Dados de registros médicos textuais incluindo 15.714 evoluções de pacientes cirúrgicos foram analisados. Depois de excluir registros vazios e que não se adequassem aos critérios do estudo, o número de registros foi reduzido para 12.637, sendo 202 evoluções de pacientes que tiveram ISC após a cirurgia e 12.435 evoluções sem ISC constatada inicialmente. Conforme visto na Tabela 4, a proporção final de registros analisados, 1,6% para 98,4%, reflete o cenário real do HCPA, de maneira semelhante as descrições cirúrgicas, sendo um conjunto de dados com classes desequilibradas.

Tabela 4. Registros textuais selecionados

Descrição	Cirurgias infectadas	Cirurgias não infectadas	Total
Número de documentos textuais disponibilizados	233	15481	15714
Registros sem texto	-2	-3037	-3039
Registros de pacientes que realizaram mais de uma cirurgia no mesmo período, das quais pelo menos uma não infectou	-3	-9	-12
Registros de infecções constatadas 30 dias após a cirurgia	-26	0	-26
Número de registros utilizados no estudo	202 (1,6%)	12435 (98,4%)	12637 (100%)

A CCIH realiza a busca por ISC metodicamente três vezes por semana. A partir de um relatório com pistas como, o uso de antibióticos, prescrição, resultados de exames microbiológicos, entre outras, é realizada a triagem das evoluções que serão lidas e discutidas para concluir se houve ou não ISC. A Tabela 5 apresenta o desempenho de cada método de classificação.

O melhor desempenho relacionado à área sob a curva ROC foi obtido pelo método Logistic Regression, considerando a relação entre verdadeiros positivos e falsos positivos. Esse resultado foi atingido utilizando 40% dos termos selecionados a partir do teste X^2 , em um conjunto de termos TF-IDF normalizados a partir da norma l1, com peso de 0,01 para a classe negativa.

Tabela 5. Desempenho dos classificadores. SA: Seleção de Atributos, P: Percentil, T: Transformação, N: Normalização, CW: Peso / Probabilidade Prévia, DP: Desvio Padrão.

Método	SA	P	T	N	CW	Precisão		Sensibilidade		ROC-AUC					
						Positivos	Negativos	Positivos	Negativos	Média	DP				
Random Forest	X ²	20%	TF	I2	1/0,01	4,5%	0,5%	99,5%	0,2%	79,3%	8,7%	73%	2,5%	76,1%	3,4%
Logistic Regression	X ²	40%	TF-IDF	I1	1/0,01	7,9%	0,8%	99,5%	0,1%	75,7%	5,4%	85,5%	1,9%	80,6%	2,4%
LinearSVC	X ²	45%	TF-IDF	I1	1/0,01	5,2%	0,4%	99,6%	0,1%	79,7%	4,1%	76,4%	1,4%	78,1%	2,5%
SVC	F-escore	25%	TF	I2	1/0,1	3,9%	2,1%	98,8%	0,2%	41,5%	9,6%	80,5%	6,8%	61%	6,3%
Nearest Centroid	X ²	80%	TF-IDF	I1	-	4,5%	0,3%	99,6%	0,1%	80,8%	5,3%	72,1%	1,1%	76,4%	2,9%
SGD	F-escore	55%	TF	-	1/0,05	2,2%	0%	100%	0%	100%	0%	27,1%	2,2%	63,6%	1%
Multinomial Naive Bayes	F-escore	80%	TF-IDF	-	20/80	22,6%	9,2%	98,9%	0,2%	29,8%	12,9%	98,3%	0,4%	64,1%	6,5%

A estratégia de classificação e pré-processamento do método Logistic Regression atingiu uma média de 80,6% de ROC-AUC (DP 2,4%), Sensibilidade média de 75,7% (DP 5,4%) para classes positivas e 85,5% (DP 1,5%) para classes negativas. Dado o desequilíbrio entre as classes, os valores médios de precisão obtidos foram de 7,9% (DP 0,8%) para a classe positiva e 99,5% (DP 0,1%) para a classe negativa. A Figura 9 apresenta a curva ROC de cada método.

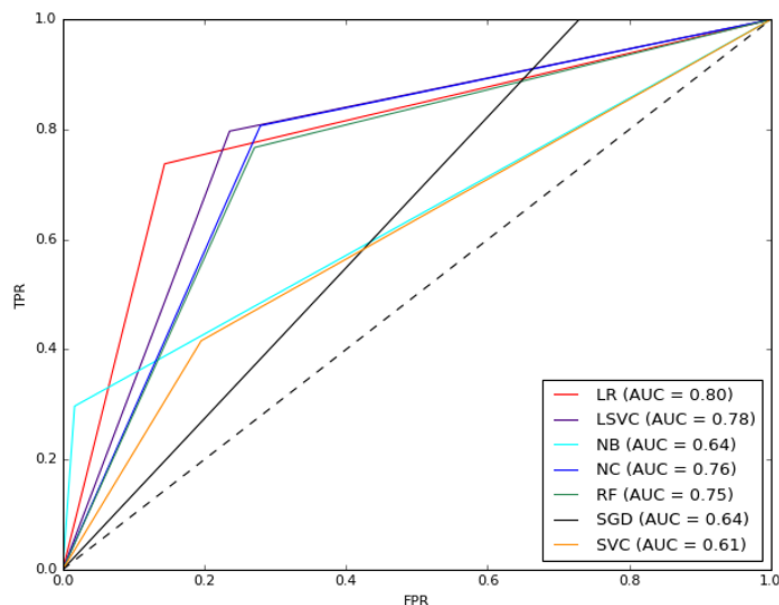


Figura 9. ROC-AUC. LSVC: LinearSVC, LR: *Logistic Regression*, NB: MultinomialNB, NC: *Nearest Centroid*, RF: *Random Forest*, SGD: *Stochastic Gradient Descent*, SVC: *Support Vector Classification*.

As estratégias de pré-processamento de cada método foram determinadas a partir de uma grade com diversas opções de parâmetros, na qual os atributos SA, P,

T, N e CW foram selecionados por terem obtido a melhor ROC-AUC. A busca pelos melhores parâmetros foi realizada por meio de uma busca aleatória em função do seu desempenho e custo computacional (BERGSTRA e BENGIO, 2012).

Os métodos de mineração de textos apresentaram potencial para reduzir o volume de evoluções a serem analisadas manualmente, exercendo um papel de triagem e direcionamento para a vigilância, conforme métodos de rastreamento de EA têm sugerido na literatura (WATCHER, 2013). A Figura 10 apresenta a proporção de Falsos Negativos e a redução do volume total de evoluções a serem analisadas a partir da utilização de cada método.

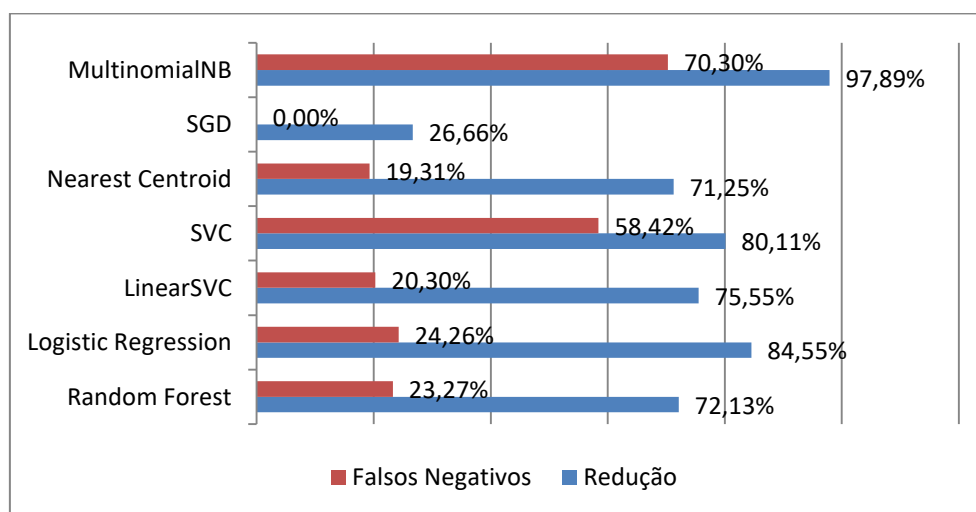


Figura 10. Redução do volume de dados vs. Falsos Negativos.

A partir dos métodos de MT utilizados, é possível reduzir 26,66% do volume de evoluções sem que infecções existentes não sejam constatadas (FN) a partir do método SGD, ou reduzir 84,55% do volume de dados tolerando 24,26% de infecções que não serão detectadas por meio do método Logistic Regression. As 202 evoluções positivas para ISC analisadas são baseadas na vigilância realizada pela CCIH, feita a partir de critérios que apontam a ocorrência de infecção. Esses critérios direcionam o trabalho de vigilância, mas podem não detectar todas as infecções existentes, dessa forma, podem haver registros verdadeiramente infectados classificados como falsos positivos.

A redução do número de registros a serem analisados manualmente apresenta um ganho para a vigilância de ISC em termos de redução de esforço e custo, porém o impacto das infecções não constatadas (FN) pelos métodos deve ser considerado. Uma maior sensibilidade positiva pode ser obtida em determinados

algoritmos, como é possível perceber no método SGD que não obteve nenhum valor FN, no entanto o número de FP, que conseqüentemente serão analisados manualmente, será maior.

O método Logistic Regression, com as opções de pré-processamento descritas anteriormente, foi escolhido como o melhor método para detecção de infecções no texto de evoluções após cirurgias (baseado na ROC-AUC). Contudo, ao buscar a redução do volume de prontuários minimizando o número de falsos negativos (infecções não constatadas), o método SGD deve ser considerado. Outras combinações de métodos e pré-processamento para detecção resultaram em escores ROC-AUC que não foram estatisticamente diferentes desses dois métodos, dado o alto desvio padrão em função do desequilíbrio das classes.

4.5 CONSIDERAÇÕES FINAIS

A mineração de textos pode ser usada no direcionamento e triagem de evoluções de pacientes cirúrgicos proporcionando redução no tempo e esforço de vigilância, apresentando potencial para aumentar o número de ISC existentes que hoje não são constatadas. O presente estudo aponta métodos eficazes para apoiar equipes de vigilância na detecção de infecções no período pós-operatório, proporcionando a melhoria da segurança do paciente. Até onde a revisão de literatura indica, este é um dos poucos estudos que utiliza métodos de MT na detecção de ISC especificamente (CAMPILLO-GIMENEZ *et al.*, 2013; MICHELSON *et al.*, 2014).

A partir dos métodos de MT utilizados, é possível reduzir 26,66% do volume de evoluções sem que infecções existentes não sejam constatadas (FN), ou reduzir 84,55% do volume de dados tolerando 24,26% de infecções que não serão detectadas. A redução do número de registros a serem analisados manualmente apresenta um ganho para a vigilância de ISC em termos de redução do esforço. Alguns algoritmos podem ter maior sensibilidade positiva garantindo que infecções não venham a ser constatadas, porém, o número de FP que conseqüentemente serão analisados manualmente, será maior.

Alguns fatores que diminuem a precisão e sensibilidade dos métodos também devem ser destacados. Evoluções de pacientes com ISC comumente possuem palavras específicas como infecção, secreção, mas o mesmo, na forma de negação,

pode não ocorrer na mesma proporção em evoluções de pacientes que não tiveram ISC, por exemplo os bigramas: *sem_infecção* e *sem_secreção*. Sendo assim, termos que semanticamente poderiam caracterizar uma não-infecção podem não ser considerados determinantes para a não ocorrência de ISC, contribuindo assim para o aumento na taxa de falsos positivos.

Outro fator relacionado ao número de falsos positivos diz respeito a narrativa médica descrita na evolução, na qual é alertado ao paciente o risco de ISC. A partir da abordagem de saco de palavras, esse alerta pode confundir classificadores de maneira que registros sem infecção possuam termos que geralmente apontam a ocorrência de infecção. Nesse cenário, a mineração de textos também pode apontar orientações para a melhoria qualitativa da narrativa médica, que venham a contribuir para que melhores resultados venham a ser alcançados a partir da MT.

O presente estudo aponta métodos eficazes para apoiar equipes de vigilância na detecção de infecções no período pós-operatório, e podem ser utilizados na detecção de outros tipos de EA, proporcionando a melhoria da segurança do paciente. Esta pesquisa buscou, de maneira exploratória, avaliar o desempenho de diferentes métodos de pré-processamento e de MT para detectar infecções em pacientes cirúrgicos.

Como trabalhos futuros, sugere-se a expansão da análise para a utilização de rótulos de classes gramaticais das palavras (*part-of-speech tagging*) e aplicação dos métodos apenas em determinados grupos de pacientes selecionados por condição médica ou faixa etária por exemplo. Sugere-se também a realização da análise de sentenças em uma análise qualitativa do texto.

4.6 REFERÊNCIAS

AGGARWAL, C. C., ZHAI, C. Mining Text Data. Springer Science Business Media, 2012.

ANVISA. Disponível em
<http://www.anvisa.gov.br/servicosaude/controle/roteiro.htm#2>. Acessado em:
 27/01/2017.

ARAMAKI, E. MIURA, Y. TONOIKE, M. OHKUMA, T. MASUICHI, H. WAKI, K. OHE, K. Extraction of adverse drug effects from clinical records. Studies in Health Technology and Informatics, 2010.

BERGSTRA, J., BENGIO, Y. Random search for hyper-parameter optimization. Journal of Machine Learning Research, 2012.

BIAN, J.; TOPALOGLU, U.; YU, F. Towards large- scale twitter mining for drug- related adverse events, 2012.

BIRD, S., KLEIN, E., LOPER, E. Natural Language Processing with Python, O'Reilly Media, 2009.

BOTSIS, T., NGUYEN, M. D., WOO, E. J., MARKATOU, M., BALL, R. Text mining for the Vaccine Adverse Event Reporting System: medical text classification using informative feature selection. Journal of the American Medical Informatics Association, 2011.

BOUZBID, S., GICQUEL, Q., GERBIER, S., CHOMARAT, M., PRADAT, E., FABRY, J., LEPAPE, A., METZGER, M-H. Automated detection of nosocomial infections: evaluation of different strategies in an intensive care unit 2000 – 2006. Journal of Hospital Infection, 2011.

BRUIJN, B. CHERRY C, KIRITCHENKO S, MARTIN J, ZHU X. Machine- learned solutions for three stages of clinical information extraction: the state of the art at i2b2 2010. Journal of the American Medical Informatics Association, 2011.

CAMPILLO-GIMENEZ, B., GARCELON, N., JARNO, P., CHAPPLAIN, J. M., CUGGIA, M. Full- text automated detection of surgical site infections secondary to neurosurgery in Rennes, France. Studies in health technology and informatics, 2013.

CHEE, B. W.; BERLIN, R.; SCHATZ, B. Predicting adverse drug events from personal health messages. AMIA Annual Symposium proceedings, 2011.

CULVER, D. H., HORAN, T. C., GAYNES, R. P., MARTONE, W. J., JARVIS, W. R., EMORI, T. G., BANERJEE, S. N., EDWARDS, J. R., TOLSON, J. S., HENDERSON, R. S., *et al.* Surgical wound infection rates by wound class, operative procedure, and patient risk index. National Nosocomial Infections Surveillance System. American Journal of Medicine, 1991.

DALTOÉ, T., BREIER, A., SANTOS, H. B., WAGNER, M. B., KUCHENBECKER, R. S. Serviços de Controle de Infecção Hospitalar: características, dimensionamento e atividades realizadas. *Revista da Sociedade Brasileira de Clínica Médica*, 2014.

DEMNER-FUSHMAN, D. CHAPMAN, W.W. MCDONALD, C.J. What can natural language processing do for clinical decision support? *Journal of Biomedical Informatics*, 2009.

DONG, Y. X., LI, Y. X., LI, J. H., ZHAO, H. Analysis on weighted AUC for imbalanced data learning through isometrics. *Journal of Computational Information*, 2012.

DUDA, R. O., HART, P. E., STORK, D. G. *Pattern Classification*. 2nd Edition. John Wiley & Sons, 2012.

SEBASTIANI, F. Machine learning in automated text categorization. *ACM Computing Surveys*, 2002.

FAYYAD, U.; PIATETSKY-SHAPIRO, G; SMYTH, P. From data mining to knowledge discovery: An overview. In *Advances in Knowledge Discovery and Data Mining*, 1996.

FELDMANN, R., SANGER, J. *The text mining handbook: Advanced approaches in analyzing unstructured data*, Cambridge Press, New York, 2006.

FREIFELD, C. C., BROWNSTEIN, J. S., MENONE, C. M., BAO, W., FILICE, R., KASSHOUT, T., DASGUPTA, N. Digital Drug Safety Surveillance: Monitoring Pharmaceutical Products in Twitter. *Drug Safety*, 2014.

FREEMAN, R., MOORE, L. S., ÁLVAREZ, L. G., CHARLETT, A., & HOLMES, A. Advances in electronic surveillance for healthcare-associated infections in the 21st Century: a systematic review. *Journal of Hospital Infection*, 2013.

FRIEDMAN, C., ELHADAD, N. *Natural Language Processing in Health Care and Biomedicine*, 2014.

GARNER, J. S. JARVIS, W.R. EMORI, T.G. HORAN, T.C. HUGHES, J.M. CDC definitions for nosocomial infections. *American Journal of Infection Control*, 1988.

GAYNES, R. P., CULVER, D. H. HORAN, T. C, *et al.* Surgical site infection (SSI) rates in the United States, 1992-1998: the national nosocomial infections surveillance system basic SSI risk index, 2001.

GENKIN, A.; LEWIS, D. D.; MADIGAN, D. Large-Scale Bayesian LogisticRegression for Text Categorization. *Technometrics*, 2007.

GRIFFIN, F.A. RESAR, R.K. *IHI Global Trigger Tool for measuring adverse events* (2nd edition). IHI Innovation Series white paper, 2009.

GURULINGAPPA, H., RAJPUT, A. M., ROBERTS, A., FLUCK, J., HOFMANN-APITIUS, M., TOLDO, L. Development of a benchmark corpus to support the automatic extraction of drug-related adverse effects from medical case reports. *Journal of Biomedical Informatics*, 2012.

GUYON, I., GUNN, S., NIKRAVESH, M., ZADEH L. A. Feature Extraction Foundations and Applications. Springer Berlin Heidelberg New York, 2006.

HAN, J. PEI, J., KAMBER, M. Data Mining: concepts and techniques - 2nd edition. Morgan Kaufmann. 2006.

HUR, J., ÖZGÜR, A., XIANG, Z., HE, Y. Identification of fever and vaccine-associated gene interaction networks using ontology-based literature mining. Journal of Biomedical Semantics, 2012.

JAPKOWICZ. N. ASSESSMENT METRICS FOR IMBALANCED LEARNING. Em: HE, H., MA. Y. Imbalanced Learning: Foundations, Algorithms, and Applications. Wiley, 2013.

KANG, N. AFZAL, Z. SINGH, B. MULLIGEN, E. M. V. KORS, J. A. Using an ensemble system to improve concept extraction from clinical records. Journal of Biomedical Informatics. Elsevier, 2012.

LUCINI, F. R., FOGLIATTO, F. S. , SILVEIRA, G. J. C., NEYELOFF, J., ANZANELLO, M. J., KUCHENBECKER, R. S., SCHAAN, B. D. Text Mining Approach To Predict Hospital Admissions Using Early Medical Records From The Emergency Department. International Journal of Medical Informatics, 2017.

MANGRAM, A. J., HORA, T. C., PEARSON, M. L., SILVER, L. C., JARVIS, W. R. Guideline for prevention of surgical infection, 1999. Infect Infection Control & Hospital Epidemiology, 1999.

MANNING, C. D., RAGHAVAN, P., SCHÜTZE, H. An Introduction to Information Retrieval. Cambridge UP, 2009.

MCCART, J. A. BERNDT, D. J. JARMAN, J. FINCH, D. K. LUTHER, S. L. Finding falls in ambulatory care clinical documents using statistical text mining. Journal of American Medical Informatics Association, 2013.

MEYSTRE, S.M. SAVOVA, G.K. KIPPER-SCHULER, K.C. HURDLE, J.F. Extracting information from textual documents in the electronic health record: a review of recent research. Yearbook Medical Informatics, 2008.

MICHELSON, J. D., PARISEAU, J. S., PAGANELLI, W. C. Assessing surgical site infection risk factors using electronic medical records and text mining. American journal of infection control, 2014.

ONG, M. S.; MAGRABI, F.; COIERA, E. Automated identification of extreme-risk events in clinical incident reports. Journal of the American Medical Informatics Association, 2012.

ORENGO, V. M., HUYCK, C. A Stemming Algorithm for the Portuguese Language. In Proceedings of the SPIRE Conference, Laguna de San Raphael, Chile, 2001.

PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., MICHEL, V., THIRION, B., GRISEL, O., BLONDEL, M., PRETTENHOFER, P., WEISS, R., DUBOURG, V., VANDERPLAS, J., PASSOS, A., COURNAPEAU, D., BRUCHER, M., PERROT, M., DUCHESNAY, E. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 2011.

PENZ, J.F.E., WILCOX, A.B., HURDLE, J.F. Automated identification of adverse events related to central venous catheters. Journal of Biomedical Informatics, 2007.

PERKINS, J. Python 3 Text Processing with NLTK 3 Cookbook. Packt Publishing. Birmingham, UK. 2014.

POSTGRESQL. Disponível em: <https://www.postgresql.org/>. Acessado em: 28/01/2017.

PYTHON SOFTWARE FOUNDATION. Disponível em: <https://www.python.org/>. Acessado em: 29/01/2017.

RAMESH, B. P., BELKNAP, S. M., LI, Z., FRID, N., WEST, D. P., YU, H. Automatically Recognizing Medication and Adverse Event Information From Food and Drug Administration's Adverse Event Reporting System Narratives. JMIR medical informatics, 2014.

ROCHEFORT, C. M., VERMA, A. D., EGUALE, T., LEE, T. C., BUCKERIDGE, D. L. A novel method of adverse event detection can accurately identify venous thromboembolisms (VTEs) from narrative electronic health record data. Journal of the American Medical Informatics Association : JAMIA, 2015.

ROSS, M. K., LIN, K.W., TRUONG, K., KUMAR, A., CONWAY, M. Text Categorization of Heart, Lung, and Blood Studies in the Database of Genotypes and Phenotypes (dbGaP) Utilizing n-grams and Metadata Features. Biomedical Informatics Insights, 2013.

SANTOS, A. C. R. B. D. MOTA, D. M. GOMES, S. M. Investigação de eventos adversos em serviços de saúde. (1ª edição). Agência Nacional de Vigilância Sanitária, 2013.

SHAREK, P. J., PARRY, G., GOLDMANN, D., BONES, K., HACKBARTH, A., RESAR, R., GRIFFIN, F. A., RHODA, D., MURPHY, C., LANDRIGAN, C. P. Performance characteristics of a methodology to quantify adverse events over time in hospitalized patients. Health services research, 2011.

SKLEARN. Disponível em: http://scikit-learn.org/stable/modules/feature_extraction.html. Acessado em: 03/02/2017 a.

SKLEARN. Disponível em: http://scikit-learn.org/stable/modules/cross_validation.html. Acessado em 04/02/2017 b.

SKLEARN. Disponível em: http://scikit-learn.org/stable/modules/grid_search.html. Acessado em 04 de janeiro de 2017 c.

SCIPY. Disponível em:
https://docs.scipy.org/doc/scipy/reference/generated/scipy.sparse.coo_matrix.html.
Acessado em: 02/02/2017.

SILVA, A. CORTEZ, P. SANTOS, M. F. GOMES, L. NEVES, J. Rating organ failure via adverse events using data mining in the intensive care unit. *Artificial Intelligence in Medicine*, 2008.

SILVA, D. A., CATEN, C. S., SANTOS, R. P., PIRES, M. R., VAZ, T. A., SANTOS, H. B. Previsão de Infecções do Sítio Cirúrgico por meio da mineração de textos no Hospital de Clínicas de Porto Alegre. Dissertação de mestrado acadêmico. Universidade Federal do Rio Grande do Sul, 2017.

TAN, A. Text Mining: The state of the art and the challenges. *Proceedings of the PAKDD 1999 Workshop on Knowledge Discovery from Advanced Databases*, 1999.

WITTEN, I. H., FRANK, E., HALL, M. A. *Data Mining: Practical Machine Learning Tools and Techniques*. Third Edition, Morgan Kaufmann, 2011.

YANG, M.; KIANG, M.; SHANG, W. Filtering big data from social media – Building an early warning system for adverse drug reactions. *Journal of Biomedical Informatics*, 2015.

YI, B.K., FALOUTSOS, C. Fast Time Sequence Indexing for Arbitrary Lp Norms. *The International Journal on Very Large Data Bases*, 2000.

ZAFARANI, R., ABBASI, M. A., LIU, H. *Social Media Mining, An Introduction*. Cambridge University Press, 2014.

5. CONSIDERAÇÕES FINAIS

Esta dissertação teve como tema a segurança do paciente, abordando o uso de métodos de mineração de textos no contexto de um hospital universitário brasileiro. O objetivo geral deste trabalho foi analisar o desempenho de métodos de mineração de textos na previsão e detecção de eventos adversos. As aplicações de diferentes estratégias de pré-processamento e MT permitiram verificar que a MT pode ser utilizada para apoiar a previsão e detecção de EA.

Buscando atingir os objetivos específicos, esta dissertação identificou as áreas da saúde e métodos de mineração de textos utilizados na prevenção e detecção de EA a partir de uma revisão sistemática da literatura. Das 54 publicações analisadas, 38 estão relacionadas a medicamentos, incluindo pesquisas sobre vacinas, reações e efeitos adversos de interações medicamentosas. Outras 16 pesquisas abordam EA de maneira geral ou específica, como infecções nosocomiais, infecções do sítio cirúrgico, quedas e tromboembolismo venoso.

A revisão sistemática da literatura apontou predominantemente para a utilização de mais de um método ou técnica de mineração de textos na mesma solução. Ao todo, 125 métodos, técnicas ou sistemas foram utilizados em 54 publicações, dos quais 79% dizem respeito a novas soluções propostas pelos autores, sendo que 16% correspondem aos métodos Support Vector Machine (SVM), classificadores Naive Bayes e Association Rule Mining e 5% aos métodos de Decision Tree, Nearest Neighbor, Logistic Regression e Clustering.

Já na avaliação do desempenho de métodos de MT quanto ao seu potencial para prever ISC, foi constatado que a MT pode ser uma ferramenta eficaz para apoiar a prevenção de infecções e possivelmente reduzir o número de ISC no período pós-operatório. A partir dos métodos de MT avaliados, foi percebido que é possível direcionar os esforços de prevenção reduzindo 55,59% do volume de cirurgias a serem acompanhadas preventivamente, com 4,26% de infecções não previstas a partir do método Logistic Regression. No mesmo sentido também é possível reduzir o volume de cirurgias a serem acompanhadas em 68,98% aceitando que 10,64% das infecções possam não ser previstas e evitadas pelo método SGD.

Por fim, a MT pode ser usada no direcionamento e triagem de evoluções de pacientes cirúrgicos com o objetivo de detectar ISC pós-operatórias. Métodos de MT proporcionam a redução no número de evoluções e também no tempo e esforço de

vigilância, apresentando potencial para aumentar o número de ISC existentes que hoje não são constatadas. A partir dos métodos SGD é possível reduzir 26,66% do volume de evoluções sem a ocorrência de falsos negativos, ou, reduzir 84,55% do volume de dados tolerando 24,26% de falsos negativos pelo método Logistic Regression.

Fatores que diminuem a precisão e sensibilidade dos métodos também foram apresentados. Um dos principais desafios nas tarefas de classificação foi o desbalanceamento das classes positivas e negativas. Em função do alto número de casos negativos para ISC, mesmo com a seleção de atributos, transformação TF-IDF e análise de bigramas e trigramas, termos como `sem_infeccao` e `sem_secrecao` não foram identificados como discriminantes na classificação dos registros sem infecção, contribuindo para o aumento na taxa de falsos positivos. A MT também pode direcionar ações para a melhoria qualitativa da narrativa médica, que venham a contribuir para que melhores resultados venham a ser alcançados pela própria mineração textual.

O presente estudo aponta métodos eficazes para apoiar equipes de vigilância na previsão e detecção de EA no período pós-operatório, proporcionando a melhoria dos cuidados e segurança do paciente. Esta pesquisa buscou, a partir da literatura e aplicações práticas, analisar diferentes métodos de pré-processamento e MT para prever e detectar EA.

Como trabalhos futuros sugere-se a expansão da análise para a utilização de rótulos de classes gramaticas das palavras (*part-of-speech tagging*) e aplicação dos métodos apenas em determinados grupos de pacientes selecionados por condição médica ou faixa etária, por exemplo, além da utilização de outras informações disponíveis no sistema informatizado do hospital. A previsão e detecção de infecções também podem ser ampliadas para outros tipos de eventos adversos. Técnicas de classificação podem ser utilizadas em documentos como prontuários, resultados de exames e prescrições de medicamentos, ampliando o número de eventos adversos previstos e detectados.