

MINERAÇÃO DE DADOS NA SAÚDE: DOENÇAS COM MAIOR INCIDÊNCIA NA POPULAÇÃO DO VALE DO AÇU/RN

Camila Barbosa; Clayton Costa

1 Instituto Federal de Educação Ciências e Tecnologia do Rio Grande do Norte,
camila.b.barbosa@outlook.com

2 Instituto Federal de Educação Ciências e Tecnologia do Rio Grande do Norte,
clayton.maciел@ifrn.edu.br

Introdução

Mineração de dados ou *data mining* surgiu em 1989 e refere-se à mineração ou descoberta de novas informações em termos de padrões ou regras com base em grande quantidade de dados. Para ser útil na prática, a mineração de dados precisa ser executada de modo eficiente em grandes arquivos e banco de dados. Embora alguns recursos estejam sendo fornecidos em SGBDRs. Portanto, nas últimas décadas, tem aumentado a necessidade de um processo automatizado para a descoberta de informações interessantes envolvendo grandes volumes de dados, em relatórios como o popular *Gartner Report*, a mineração de dados foi definida a área mais ampla, chamada *descoberta de conhecimento* (Elmasri; Navathe 2012, p.698). Assim, a descoberta de conhecimento em bases de dados, ou *knowledge Discovery on databases* (KDD), consiste na atividade de transformar grandes quantidades de dados em informação que possa ser usada na prática, ou que produza relevância em determinada área de conhecimento.

Pesquisadores motivados pelo desafio de transformar informação em conhecimento, logo se deparam com o *Knowledge Discovery in Databases* (KDD), dando ênfase na aplicação de mineração de dados (12ª CONTECSI, 2015). No campo da saúde, existem diversos trabalhos realizados com a aplicação da mineração de dados envolvendo diversas informações, dentre eles o trabalho de *identificação de Padrões em Registros de Doenças com Técnicas de Mineração de Dados*, que tem como evidência aplicar as técnicas de mineração de dados para descobrir em bases de dados da área da saúde, paciente com Doença de Chagas, com o intuito de classificar as principais características desses indivíduos. E *Mineração de dados em triagem de risco de saúde* onde seu objetivo é descrever, passo a passo, a aplicação de um processo de descoberta de conhecimento em banco de dados (KDD) no domínio da triagem médica.

Dessa forma, considerando essa problemática de ordem pessoal, social e profissional que convergem para uma articulação entre a área da

informática e da saúde. Sabe-se que essas duas áreas de conhecimento estuda a aplicação dos recursos tecnológicos no ensino, na prática, na assistência e gerenciamento da assistência, onde essa relação se tornou cada vez mais importante para a saúde no Brasil. O motivo da escolha desse tema se torna um grande desafio de se alcançar, pois fazer uma extração de dados de prontuários médicos é de muito sigilo e muita discricção, havendo ética, seriedade e respeito, principalmente com os pacientes do Hospital. Este trabalho tem como objetivo determinar as doenças prevalentes na população do Vale do Açu/RN por meio do processo ETL e de técnicas de mineração de dados. E como objetivos específicos aplicar o processo ETL para a extração de dados das fichas de atendimento, identificando as principais doenças que foram acometidas na população, investigando os fatores que contribuem para a incidência de doenças no Vale do Assú/RN.

Metodologia

O primeiro passo para a realização deste trabalho consiste na investigação bibliográfica, pautada na leitura de trabalhos que trata de mineração de dados com ênfase na saúde, e como a mineração pode se aplicada na saúde, bem como também a importância das contribuições da mineração para a saúde. Após as investigações bibliográficas será feita uma coleta de dados a partir das fichas de atendimento dos pacientes que foram internados no Hospital Regional Nelson Inácio dos Santos localizado na cidade de Assú/RN, onde esses dados serão inseridos no programa Excel que é software livre e então será aplicado o processo ETL (Extração, Transformação e Carregamento) para extrair e transformar os dados obtidos dos prontuários médicos em dados legíveis em software de mineração.

Para a validação da proposta deste trabalho será utilizado o WEKA que é um software gratuito e de código aberto que pode ser usado para minerar dados e transformar em novo conhecimento. E assim o objetivo deste trabalho será concretizado, que é identificar e analisar quais doenças foram prevalentes na população do Vale do Assú/RN, ocorridas em determinadas datas ou estações de ano.

Resultados e discussão

A partir das informações obtidas da coleta dos dados dos prontuários dos pacientes, e assim esses dados poderão ser extraídos e minerados para identificar e analisar quais doenças é prevalente em um determinado ano ou mês em que foram registradas e notificadas no Hospital Regional Nelson Inácio dos Santos. Assim, espera-se que o trabalho venha contribuir de forma significativa para a saúde da população na região do Vale do Assú através de medidas de promoção e prevenção da saúde. Bem como, auxiliar a administração do hospital e a equipe médica

sobre como determinadas doenças vem sendo acometidas na população.

Conclusões

Portanto, o trabalho só foi possível de ser realizado através de uma pesquisa bibliográfica com base em trabalhos de alguns autores, onde se pôde obter um vasto conhecimento e novas informações importantes, e assim conseguir as fichas médicas para obter as informações necessárias para a validação deste trabalho. Entende-se que o problema a ser tratado em KDD (descoberta de conhecimento em bases de dados) é o de transformar um conjunto de dados em um modelo utilizável, com determinado propósito, e que possibilite o entendimento humano. Um exemplo seria a possibilidade de prever a ocorrência de uma importante situação em determinado domínio de conhecimento, como a descoberta de doenças que foram mais prevalentes em determinada data. O campo da saúde, é uma das áreas de maior importância para a aplicação de mineração de dados, podendo potencializar o auxílio no controle de infecções, diagnósticos e tratamento de várias doenças, além de gestão de recursos da saúde, gestão hospitalar e administração da saúde pública.

Palavras-Chave: Pesquisa bibliográfica; Descoberta de conhecimento; Saúde.

Fomento

Instituto Federal de Educação Ciências e Tecnologia do Rio Grande do Norte.

Referências

ELMASRI, Ramez, SHAMKANT, B. Navathe. **Sistemas de banco de dados**. ed. 6º, p.1-766, São Paulo, ag. 2012.

EDUCAÇÃO, Colunista Portal - et al. **O QUE É SAÚDE?** 2013. Disponível em: <<https://www.portaleducacao.com.br/direito/artigos/32304/o-que-e-saude>>. Acesso em: 7 nov.2016.

SOUZA, Jovani Taveira de et al. **Criação de Conhecimento através de Data Mining:** Um estudo sobre regras de associação em uma base de dados do Weka. 2015. Disponível em: <<http://www.revistaespacios.com/a16v37n06/16370624.html>>. Acesso em: 03 jan. 2017.

MANAGEMENT, Contecsi - International Conference On Information Systems And Technology et al. **Identificação de Padrões em Registros de Doenças com Técnicas de Mineração de Dados**. 2015. Disponível em: <<http://www.contecsi.fea.usp.br/envio/index.php/contecsi/12CONTECSI>>. Acesso em: 7 nov. 2016.