

UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE INFORMÁTICA

RAFAEL TEIXEIRA SOUSA

**Predição de Complicações em Pacientes
Diabéticos Utilizando Dados
Transacionais de Operadoras de Planos
de Saúde**

Goiânia
2020

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR
VERSÕES ELETRÔNICAS DE TESES E DISSERTAÇÕES
NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das dissertações e teses disponibilizados são de responsabilidade exclusiva dos autores. Ao encaminhar(em) o produto final, o autor e o orientador firmam o compromisso de que ele não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do material bibliográfico: [] Dissertação [X] Tese

2. Identificação da Tese ou Dissertação:

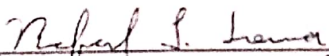
Nome completo do autor: Rafael Teixeira Sousa

Título do trabalho: Predição de complicações em pacientes diabéticos utilizando dados transacionais de operadoras de planos de saúde

3. Informações de acesso ao documento:

Concorda com a liberação total do documento [] SIM [X] NÃO¹

Independente da concordância com a disponibilização eletrônica, é imprescindível o envio do(s) arquivo(s) em formato digital PDF da tese ou dissertação.


Assinatura do(a) autor(a)²

Ciente e de acordo:


Assinatura do(a) orientador(a)²

Data: 23/04/2020

¹ Neste caso o documento será embargado por até um ano a partir da data de defesa. O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

² As assinaturas devem ser originais sendo assinadas no próprio documento, imagens coladas não serão aceitas.

RAFAEL TEIXEIRA SOUSA

Predição de Complicações em Pacientes Diabéticos Utilizando Dados Transacionais de Operadoras de Planos de Saúde

Tese apresentada ao Programa de Pós-Graduação do Instituto de Informática da Universidade Federal de Goiás, como requisito parcial para obtenção do título de Doutor em Ciência da Computação.

Área de concentração: Ciência da Computação.

Orientador: Prof. Anderson da Silva Soares

Goiânia
2020

Resumo

Sousa, Rafael Teixeira. **Predição de Complicações em Pacientes Diabéticos Utilizando Dados Transacionais de Operadoras de Planos de Saúde**. Goiânia, 2020. 96p. Tese de Doutorado

Instituto de Informática, Universidade Federal de Goiás.

O avanço da taxa de mortalidade e custos relacionados as doenças crônicas tem aumentado significativamente em uma escala global. São doenças cujo tratamento inclui mudanças de hábitos e monitoramento constante por profissionais de saúde. A avaliação de risco em pacientes crônicos é um elemento importante no cuidado preventivo, pois identifica pacientes de alto risco futuro. A Inteligência Artificial vem sendo usada como ferramenta para esse tipo de avaliação por meio de dados clínicos de Prontuário Eletrônico do Paciente (PEP). Devido a baixa adesão, falta de interoperabilidade e dificuldade de acesso ao PEP este trabalho propõe o uso de dados transacionais de operadoras de planos de saúde como alternativa para avaliação de crônicos. Para viabilizar a modelagem por meio desses dados este trabalho apresenta alguns métodos de aprendizado de representações com redes neurais profundas (Deep Learning). Os métodos propostos são: uma representação vetorial densa para códigos TUSS capaz de capturar relações e significados; um método de representação de informação de datas em forma vetorial a fim de levar à compreensão temporal a quaisquer modelos de Machine Learning e uma arquitetura de rede neural com capacidade de processar sequências de dados transacionais para reconhecer padrões. Toda a proposta foi avaliada em um estudo de caso na modelagem da diabetes para a predição de complicações relacionadas a doença em duas operadoras. O desempenho obtido foi uma AUC de 0,92 na predição de complicações com seis meses de antecedência. Por fim foi realizada uma avaliação do risco calculado a fim de compreender o comportamento do modelo. Foram encontradas relação de acordo com a literatura médica como o aumento do risco com a idade. Algumas limitações foram encontradas na proposta, como a dificuldade em determinar quais indivíduos possuem condições crônicas devido à falta de informações clínicas. A proposta vem apresentando êxito na identificação de indivíduos de risco e no aprimoramento dos programas de monitoramento das operadoras, que já utilizam regularmente a solução.

Palavras-chave

Doenças Crônicas, Deep Learning, RNN

Abstract

Sousa, Rafael Teixeira. **Representation Learning on Transaction Data of Health Care Plans to Model Chronic Diseases**. Goiânia , 2020. 96p. PhD. Thesis
Instituto de Informática, Universidade Federal de Goiás.

The advance of mortality rate and costs related to chronic diseases have been increasing in a global scale. Those are diseases that treatment need change of habits and constant monitoring by healthcare professionals. The risk assessment in chronic patients is an important element on prevention care, because it can identify patients with high future risk. Artificial Intelligence have been used as a tool for this kind of evaluation through clinical data from Electronic Health Records (EHR). Due to its low adhesion, lack of interoperability e access barriers to HER this work proposes the use of transaction records from health care plans as an alternative to assess chronic patients. To make it feasible the modeling through this kind of data this work exposes some methods of representation learning with neural networks (Deep Learning). The proposed methods are: a vector representation to TUSS codes capable of capture meanings and relations; a method to represent dates in vectors to bring time comprehension to any Machine Learning model and a neural network architecture to deal with transaction data and recognize patterns. All the proposal was evaluated on a case study on modeling diabetes to predict complication on two different companies. The performance obtained was an AUC of 0.92 on the prediction of complications in six months. By the end it was evaluated the calculated risk to understand the model behavior. It was found relations that agree with medical literature as the risk increase with age. Some limitations were found on the proposal, as the difficulty to define which individuals have chronic diseases due to the lack of clinical data. The proposal have been successfully identifying individuals with high risk and improving monitoring programs of two companies, which already regularly have been using it.

Keywords

Chronic Diseases, Deep Learning, RNN.