Preservação de Privacidade em Aprendizagem de Máquina por meio de Criptografia Homomórfica

Renan Chaves Bezerra [1], **Eanes Torres Pereira** [2]

RESUMO

Com o sucesso da Image-Net, diversas áreas do conhecimento humano começaram a aplicar redes neurais profundas em outras tarefas. No entanto, imagens médicas, assim como outros tipos de dados para auxílio ao diagnóstico médico, possuem informações sensíveis que podem prejudicar os pacientes caso sejam divulgadas. Neste ponto, chega-se a um dilema: ao mesmo tempo em que são necessárias centenas de milhares de amostras de dados para treinar os modelos, raramente um único indivíduo ou entidade possui todos os dados necessários para treinar os modelos. Neste projeto, propõe-se a investigação e desenvolvimento de uma abordagem que permita treinar modelos de aprendizagem de máquina para reconhecimento de padrões utilizando criptografia homomórfica (CH). A revisão da literatura revelou que há poucos trabalhos escritos sobre a aplicação de CH em treinamento de modelos profundos. Além disso, não foi encontrada nenhuma patente em português no Brasil que tratasse desse tema. A metodologia adotada constituiu das seguintes etapas: estudo de criptografia, estudo de criptografia homomórfica, análise de abordagens e bibliotecas de softwares, implementação de CH associada ao treinamento de modelos de aprendizagem profunda para classificação de dados de fraudes em cartões de crédito e avaliação da abordagem.

Palavras-chave: inteligência artificial, privacidade, dados bancários.

[1] Aluno de Ciências da Computação, Unidade Acadêmica de Sistemas e Computação, UFCG, Campina Grande, PB, e-mail: renan.bezerra@ccc.ufcg.edu.br

[2] Doutor, Orientador, Unidade Acadêmica de Sistemas e Computação, UFCG, Campina Grande, PB, e-mail: eanes@computacao.ufcg.edu.br

Preservação de Privacidade em Aprendizagem de Máquina por meio de Criptografia Homomórfica

ABSTRACT

With the success of Image-Net, several areas of human knowledge began to apply deep neural networks in other tasks. However, medical images, as well as other types of data to aid medical diagnosis, contain sensitive information that could harm patients if disclosed. At this point, a dilemma arises: while it takes hundreds of thousands of data samples to train the models, rarely does a single individual or entity have all the data needed to train the models. In this project, we propose the investigation and development of an approach that allows training machine learning models for pattern recognition using homomorphic cryptography (HC). The literature review revealed that there are few works written about the application of CH in deep model training. In addition, no patent in Portuguese was found in Brazil dealing with this topic. The adopted methodology consisted of the following steps: study of cryptography, study of homomorphic cryptography, analysis of approaches and software libraries, implementation of HC associated with the training of deep learning models for classification of credit card fraud data and evaluation of the approach.

Keywords: artificial intelligence, privacy, bank data.

[1] Aluno de Ciências da Computação, Unidade Acadêmica de Sistemas e Computação, UFCG, Campina Grande, PB, e-mail: renan.bezerra@ccc.ufcg.edu.br

[2] Doutor, Orientador, Unidade Acadêmica de Sistemas e Computação, UFCG, Campina Grande, PB, e-mail: eanes@computacao.ufcg.edu.br