



APLICAÇÃO DE ALGORITMOS DE APRENDIZAGEM DE MÁQUINA PARA ANÁLISE DO PERFIL DEMOGRÁFICO E SOCIOECONÔMICO DO MAL DE PARKINSON NO BRASIL

Odolir Daniel dos Santos Junior, Esther Gregório, Renata de Souza, Anita Maria da Rocha Fernandes.

Engenharias e Ciências Agrárias, Exatas e da Terra
Ciência da Computação - Sistemas de Computação

Em um cenário mundial de mudanças sociodemográficas e epidemiológicas, constata-se um intenso envelhecimento populacional ao qual se associa um significativo aumento de doenças crônicas. Neste contexto, na atualidade, a doença de Parkinson tem sido reconhecida como a segunda doença neurodegenerativa mais frequente, cujas adversidades impactam indivíduos a partir de 40 anos em uma fase socialmente produtiva. É importante para os gestores de saúde entenderem o perfil demográfico e socioeconômico da doença a fim de fortalecer e/ou criar políticas públicas que amparem as pessoas com Parkinson. Alguns estudos vêm sendo feitos a respeito das questões socioeconômicas e demográficas da doença, a fim de traçar perfis sob o assunto, que respaldem os gestores na execução ou elaboração de políticas públicas. São muitos os dados que podem auxiliar na análise demográfica e socioeconômica da doença e tais dados são oriundos de diversas bases, tais como as bases de dados do Ministério da Saúde (SIH - Sistema de Internações Hospitalares e SIM - Sistema de Informações de Mortalidade) e do IBGE, e do IPEA. Para analisar esta grande quantidade de dados produzida a cada hora, os algoritmos de Aprendizado de Máquina ou Machine Learning passam a ser uma ferramenta importante. Machine Learning refere-se ao funcionamento de sistemas computacionais capazes de aprender e modificar o seu comportamento em resposta a estímulos externos, ou através de experiências acumuladas durante sua operação. A flexibilidade dos métodos de Machine Learning os torna mais apropriados para situações nas quais existe pouco conhecimento a priori sobre o domínio, e/ou esse conhecimento a priori é de difícil elicitação. Sendo assim, esta pesquisa visa contribuir para a compreensão do perfil da Doença de Parkinson, analisando os dados sob a ótica demográfica e socioeconômica, considerando a população das cidades e regiões, bem como seus PIBs, dentre outros dados. Para isto, as seguintes etapas foram seguidas: (i) estruturação da base de dados: para isto utilizou-se a base de dados do DATASUS - SIM e a base do IBGE sobre os PIBs, considerando o período de 2009 a 2019, com um total de 12.813.859 registros; (ii) limpeza e preparação dos dados, utilizando Pandas e aplicando análise descritiva dos dados, o que levou a diminuição da base para 7.815 registros; (iii) análise dos algoritmos: considerando a natureza do estudo, optou-se pelo aprendizado não-supervisionado, mais especificamente o algoritmo de K-means, o qual é um dos mais simples e mais rápidos métodos de clusterização. Ele tenta separar os dados em K (um número predefinido) clusters, de acordo com a distância de cada ponto até algo chamado de centroide. Utilizou-se a biblioteca scikit learning do Python. Percebeu-se que os estados da região Sudeste formaram um cluster de destaque, e que o estado do



Maranhão em 2014 apresentou uma anormalidade de incidência praticamente zero de parkinsonianos. Em pesquisa paralela realizada, verificou-se que provavelmente este problema tenha sido decorrente a falta de cadastro dos dados do estado para o CID referente a Parkinson no sistema do DATASUS para o referido ano. Considerando, então a região sudeste, verificou-se que as políticas públicas para Parkinson nesta região foram mais efetivas, enquanto em outras regiões do país ainda estão sendo divulgadas. Os hábitos alimentares dos parkinsonianos também foram considerados visto que a região sudeste concentra o maior PIB do país, porém este estudo ainda não foi concluído. A partir dos dados obtidos no estudo e considerando que há muitas informações a serem agregadas para analisar o Mal de Parkinson no país, como resultado deste estudo, foi proposta a criação de um Portal de Informações sobre a Doença de Parkinson.

Palavras-chave: Perfil Demográfico, Aprendizagem de Máquina, Mal de Parkinson.

Programa UNIEDU – Bolsa de Pesquisa Art. 170 e Art. 171 / Governo de Santa Catarina / UNIVALI