

UTILIZAÇÃO DE APRENDIZADO DE MÁQUINA E DADOS CLÍNICOS PARA A CLASSIFICAÇÃO DE CASOS DE LEPTOSPIROSE HUMANA

4º Encontro Nacional de Epidemiologia Veterinária, 4ª edição, de 19/07/2022 a 21/07/2022
ISBN dos Anais: 978-65-81152-81-9

BOHM; Bianca Conrad¹, SILVA; Suellen Caroline Matos², BRUHN; Fábio Raphael Pascotti³

RESUMO

A leptospirose apesar de ser uma zoonose com ampla distribuição geográfica e ser considerada um problema de saúde pública mundial, ainda é tratada de forma negligenciada. O objetivo deste trabalho é utilizar a inteligência artificial para realizar o diagnóstico (triagem) de casos de leptospirose humana utilizando dados do paciente (sexo, raça e idade) e dos sinais clínicos. Trata-se de uma pesquisa quantitativa, realizada com dados secundários provenientes das fichas de notificação individual do Sistema Nacional de Agravos de Notificação (SINAN), obtidas no site DataSUS. Foi realizado o download das fichas de notificação de leptospirose do estado do Rio Grande do Sul, no período de 2007 a 2019. As variáveis independentes selecionadas foram: sexo, raça, idade, sinais clínicos (febre, mialgia, cefaleia, prostração, congestão, dor na panturrilha, vômito, diarreia, icterícia, alteração renal e respiratória) e a variável dependente foi a classificação final (caso de leptospirose confirmado ou descartado por exame laboratorial). O teste de qui-quadrado foi realizado para avaliar a associação entre as variáveis independentes e a variável dependente, sendo que somente aquelas associadas ($p < 0,05$) com a dependente foram selecionadas para construção do modelo. Após a seleção das variáveis de interesse foi realizada a limpeza do banco, sendo retirados todos os casos em que o preenchimento da ficha não foi completo. Apenas casos confirmados ou descartados por critério laboratorial foram utilizados. Neste estudo foi utilizado o modelo Naive Bayes, disponível no software Weka (Waikato Environment for Knowledge Analysis). O modelo foi treinado utilizando a técnica de k-fold. A eficácia do modelo no diagnóstico da leptospirose foi avaliada pelas métricas de acurácia, sensibilidade e especificidade. Todas as variáveis independentes testadas se mostraram associadas a incidência da leptospirose ($p < 0,05$), portanto, foram incluídas no modelo. O banco utilizado era formado por 5.361 casos confirmados e 5.361 casos descartados. Os valores das métricas de classificação do modelo Naive Bayes foram: 63,3% de acurácia, 63% de sensibilidade e 63,9% de especificidade. A classificação dos casos de leptospirose através das informações selecionadas não foi satisfatória, visto que os valores das métricas de avaliação foram baixos. A inclusão de variáveis ligadas ao risco de exposição a leptospirose e informações socioeconômicas podem melhorar as métricas de classificação do modelo. **Agências de Fomento:** O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Os autores agradecem à FAPERGS- Fundação de Amparo a Pesquisa Do Rio Grande do Sul - pelo apoio financeiro para a realização desta pesquisa (código de financiamento 21/2551-0000608-0).

PALAVRAS-CHAVE: Doenças Negligenciadas, Inteligência Artificial, Naive Bayes, SUS, Zoonoses

¹ Programa de Pós-graduação em Veterinária - Universidade Federal de Pelotas - UFPel, biankabohm@hotmail.com

² Programa de Pós-graduação em Veterinária - Universidade Federal de Pelotas - UFPel, suellen.carol.as@gmail.com

³ Centro de Controle de Zoonoses da Universidade Federal de Pelotas - UFPel, fabio_rpb@yahoo.com.br