

11ª Jornada Científica e Tecnológica do IFSULDEMINAS & 8º Simpósio de Pós-Graduação

MACHINE LEARNING APLICADO À ANÁLISE DE ECG

Samuel A. MEZAVILA¹, Nykollas A. A. DIAS²; Matheus E. FRANCO³

RESUMO

A inserção da Inteligência Artificial (IA) nas atividades do dia-a-dia é uma realidade. Além das práticas diárias comuns, o aprendizado de máquina aplicado na medicina é algo importante, sendo uma das principais áreas de estudo da IA. Diagnósticos médicos podem ter apoio com a utilização de algoritmos inteligentes, tendo em vista a rapidez e eficiência que podem proporcionar. Este artigo apresenta um estudo sobre a implementação de um algoritmo de aprendizagem de máquina utilizando a linguagem Python para análises de eletrocardiogramas, classificando o risco do paciente de sofrer infarto do miocárdio.

Palavras-chave: Inteligência artificial; Informática aplicada à saúde; Aprendizagem de máquina.

1. INTRODUÇÃO

No campo da inteligência artificial, as máquinas inteligentes, na maior parte dos anos, foram vistas como algo fora da realidade. Ainda não vemos “carros voadores”, mas dispositivos eletrônicos com capacidade de aprendizado já são realidade.

Na medicina, vemos que a inserção da inteligência artificial é algo importante, tendo em vista que permite uma gama maior de possibilidades para tratamentos complexos. Diversas técnicas de aprendizagem de máquina foram desenvolvidas com objetivo de auxiliarem na previsão de doenças graves e ajudar no diagnóstico prematuro.

Diante da importância da discussão do referido tema na atualidade, o presente trabalho busca o estudo da aplicação de técnicas de aprendizagem de máquina para análises de exames de eletrocardiograma (ECG), a fim de auxiliar no diagnóstico do risco de infarto do miocárdio. Ressalta-se que a proposta não é um sistema médico, nem objetiva diagnosticar ou tratar quaisquer condições, tratando-se de um estudo que pode vir a auxiliar o diagnóstico cardíaco.

2. FUNDAMENTAÇÃO TEÓRICA

Segundo Winston (1992), a Inteligência Artificial (IA) é o estudo do projeto de agentes inteligentes. As primeiras pesquisas de IA nos anos 1950 entraram em temas como a resolução de problemas, aprendizagem automática e redes neurais.

1 Graduando em Sistemas de Informação, IFSULDEMINAS – Campus Machado. samuelmezavila@gmail.com.

2 Graduando em Sistemas de Informação, IFSULDEMINAS – Campus Machado. nicolauarti09@gmail.com

3 Orientador, IFSULDEMINAS – Campus Machado. matheus.franco@ifsuldeminas.edu.br.

Pesquisadores como Anderson (2007) e Obermayer (2016) preveem que ferramentas de aprendizagem de máquina e inteligência artificial para analisar grandes quantidades de dados será um requisito fundamental para a nova geração de médicos, e também que algoritmos e hardwares automatizados poderão substituir médicos algumas áreas.

A incorporação do aprendizado de máquina na medicina clínica promete melhorar substancialmente a prestação de cuidados de saúde. Empresas privadas estão correndo para construir o aprendizado de máquina na tomada de decisões médicas, seguindo ambas as ferramentas que suportam os médicos e os algoritmos projetados para funcionar independentemente deles (CHAR; SHAH; MAGNUS 2018).

A mente médica está se tornando a combinação de livros publicados na área e dados registrados nos sistemas de saúde, em oposição à experiência clínica individual de cada profissional (CHAR, SHAH e MAGNUS, 2018). Embora essa mudança apresente oportunidades estimulantes para aprender com dados agrupados, a memória eletrônica pode assumir uma importância que talvez nunca tenha sido planejada. Os médicos podem se direcionar à aprendizagem de máquina para o diagnóstico e aconselhamento sobre tratamentos, e não apenas como uma ferramenta de suporte.

3. MATERIAL E MÉTODOS

Os dados para classificação de batimentos cardíacos foram obtidos por meio do portal PhysioNet, nomeado como “MIT-BIH Arrhythmia Dataset” (MOODY e MARK, 2001), contendo aproximadamente 110.000 dados. Os dados são divididos em classes, baseados na AAMI EC57 (ANSI-AAMI, 1998): N (normal); S (contração atrial prematura); V (contração ventricular prematura); F (fusão de ventricular e normal) e Q (inclassificável).

Para processamento dos dados utilizou-se técnicas de aprendizagem de máquina (*deep* através linguagem de programação Python e de sua biblioteca scikit-learn. Os parâmetros utilizados e o código implementado está disponível em <http://bit.ly/2SUjJCo>. A metodologia de desenvolvimento do estudo baseou-se no trabalho de Kachuee, Fazeli e Sarrafzadeh (2018).

4. RESULTADOS E DISCUSSÕES

O treinamento da rede levou 2 horas e 34 minutos em um computador com 12GB de RAM, CPU com 4 núcleos de 3.9Ghz. O modelo utilizado (Figura 1) é uma Perceptron Multicamadas (MLP - *Multi Layer Perceptron*). A MLP utilizada contém 4 camadas ocultas, sendo a camada de entrada e a primeira com 187 neurônios, o que corresponde a quantidade de atributos do *dataset*, a segunda com 374, a terceira com 94 e a quarta com 42. A função de ativação para cada camada é a ReLU, amplamente utilizada em redes neurais. Consiste em uma função não linear, podendo ter várias camadas de neurônios ativadas por ela.

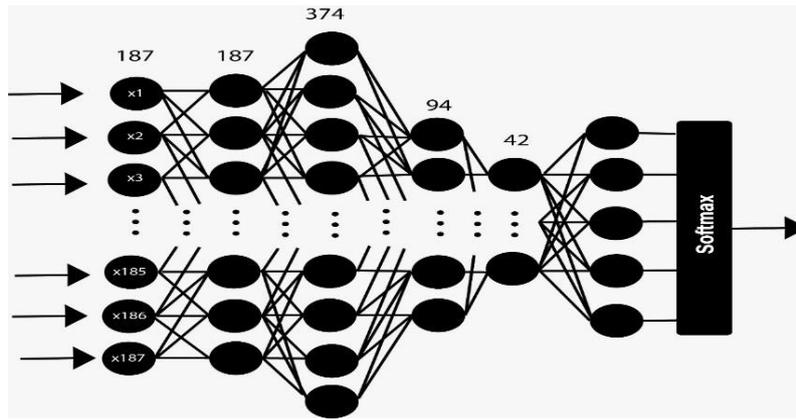


Figura 1 – Modelo de arquitetura da rede neural MLP utilizada.
 Fonte: Autoria Própria

O treinamento foi realizado em 200 épocas, que é cada iteração que o modelo realiza para aprimorar seus parâmetros. Em sua totalidade, foram separados 25% dos dados para teste e 75% para treinamento da rede.

A Figura 2 apresenta a matriz de confusão resultante da classificação. Os números nos blocos são as amostras classificadas em cada categoria, normalizadas com os valores entre 0 e 1, arredondadas para 2 dígitos. Assim, o modelo é capaz de fazer previsões precisas e distinguir as classes para cada análise.

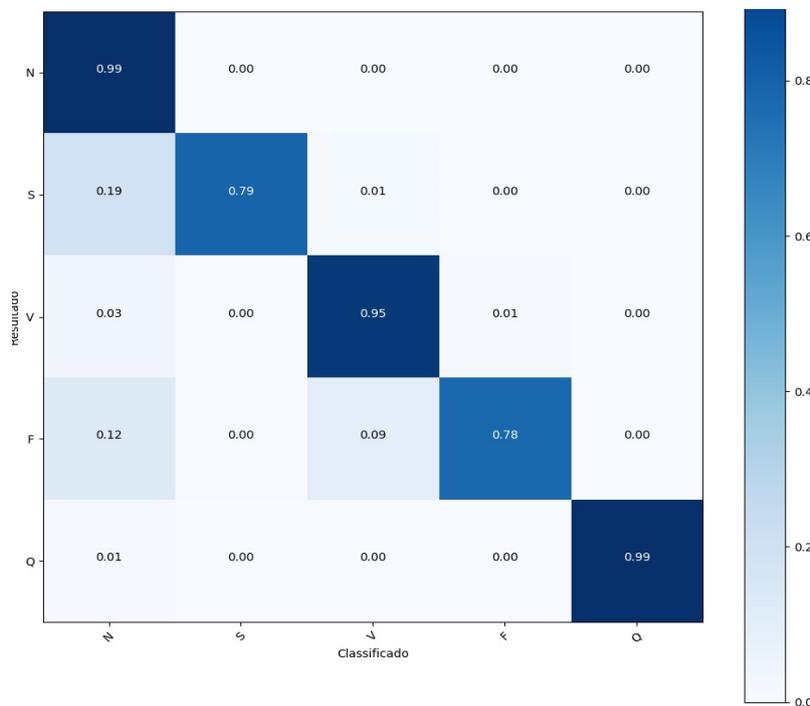


Figura 2 – Matriz de confusão sem normalização
 Fonte: Autoria própria, adaptado de Kachuee, Fazeli e Sarrafzadeh (2018).

Na figura 3A, observa-se um exemplar de cada batimento cardíaco usado para a identificação e separação das classes no treinamento da rede neural. A classificação final dos 109.446 registros ficaram dispostos conforme a figura 3B, atingindo a acurácia no modelo treinado de 0,98, algo superior ao obtido no trabalho de Kachuee, Fazeli e Sarrafzadeh (2018), o qual baseou-se este estudo.

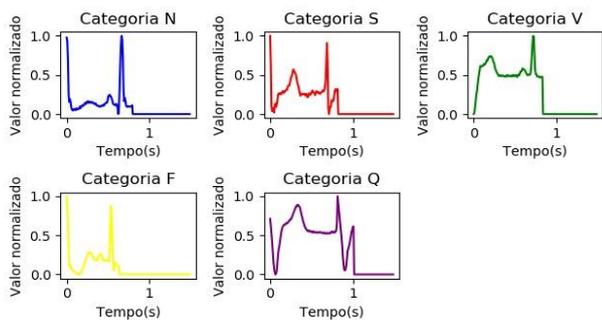


Figura 3A – Exemplos de batimentos cardíacos
Fonte: Autoria própria

Categoria	Quantidade de dados incluídos
N	90.589
S	8.039
V	7.236
F	2.779
Q	803

Figura 3B – Classificação dos dados
Fonte: Autoria própria.

5. CONCLUSÕES

A utilização das técnicas de aprendizagem de máquina aplicada à saúde é uma realidade, conseguindo analisar grandes quantidades de dados, chegando a resultados satisfatórios. Neste trabalho, buscou-se estudar e implementar técnicas para análise de dados de ECG. Com o modelo de treinamento utilizado foi possível se atingir uma acurácia de alto nível, sendo um ponto de partida para novos estudos a serem realizados. Ademais, as tecnologias e técnicas utilizadas, podem ser aplicadas para outras soluções.

REFERÊNCIAS

- ANDERSON, Susan Leigh; ANDERSON, Michael. Machine Ethics: Creating an Ethical Intelligent Agent. Association for the Advancement of Artificial Intelligence, [S. l.], p. 15-25, 10 dez. 2007.
- ANSI-AAMI. Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms. Association for the Advancement of Medical Instrumentation, Arlington, VA, 1998.
- CHAR, Danton; H. SHAH, Nigam; MAGNUS, David. Implementing Machine Learning in Health Care — Addressing Ethical Challenges. The New England Journal of Medicine, p. 1, 15 mar. 2018.
- GOLDBERGER, Ary L. et al. PhysioBank, PhysioToolkit, and PhysioNet, 2000.
- KACHUEE, Mohammad; FAZELI, Shayan; SARRAFZADEH, Majid. Ecg heartbeat classification: A deep transferable representation. In: IEEE I.C. on Healthcare Informatics (ICHI), 2018.
- KONONENKO, Igor. Machine learning for medical diagnosis: history, state of the art and perspective. Science Direct, [S. l.], p. 89-109, 2 ago. 2001.
- MOODY, G. B.; MARK, R. G. “The impact of the mit-bih arrhythmia database,” IEEE Engineering in Medicine and Biology Magazine, vol. 20, no. 3, pp. 45–50, 2001
- OBERMAYER Z.; E.J., EMANUEL. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. US National Library of Medicine, [S. l.], p. 1-8, 29 set. 2016.
- RIGBY, Michael J. Ethical Dimensions of Using Artificial Intelligence in Health Care. AMA Journal of Ethics, [S. l.], p. 1, 1 fev. 2019.