

# Universidade de Pernambuco

## Programa de Pós-Graduação em Engenharia da Computação (PPGEC)

### Proposta de Dissertação de Mestrado

**Área: Computação Inteligente**

**Título: Aplicação de Inteligência Artificial na área médica**

**Orientador – Roberta Andrade de A. Fagundes (roberta.fagundes@upe.br)**

**Co-orientador: Wylliams Barbosa Santos (wbs@upe.br)**

A medida que a demanda por assistência médica de qualidade continua cresce, o uso do aprendizado de máquina na área da saúde ganha mais espaço. Isso é possível pela qualidade e agilidade que processos guiados por inteligência artificial são capazes de entregar. O termo aprendizado de máquina (sinônimo em inglês - Machine Learning - ML), é um tipo de inteligência artificial (IA). O ML na área médica está se tornando mais amplamente usado e está ajudando pacientes e profissionais de muitas maneiras diferentes, como por exemplo, identificar relações significativas em dados brutos, decisões de tratamento, assistência ao paciente, decisões financeiras e operacionais.

Na comunidade de ML, um dos tópicos de crescente interesse é a pesquisa de novas arquiteturas de aprendizado profundo (Deep Learning) para problemas de classificação e regressão em grandes massas de dados [1]. Nesse contexto, um dos principais gargalos para a utilização de redes de aprendizado profundo é o seu elevado tempo de aprendizado. Uma possível solução para redução da complexidade é a utilização da Extreme Learning Machine (ELM) [2], onde um conjunto mínimo de parâmetros precisa ser ajustado durante o processo de aprendizado. Além disso, rapidez no treinamento, e boa capacidade de generalização são vantagens do ELM. Entretanto, uma das desvantagens é a escolha aleatória dos pesos da camada oculta, podendo ocasionar problemas de multicolinearidade e aumento da variância nas estimativas dos modelos gerados. Para isso, os modelos de regressão Ridge e LASSO [3] resolvem o problema de multicolinearidade acrescentando um pequeno vício ao estimador dos coeficientes de regressão encontrado por mínimos quadrados ordinários, afastando o sistema da singularidade. Tais modelos, apesar de viciarem o estimador, possibilitam encontrar resultados com menores erros quadráticos médios, à medida que diminui consideravelmente a variância dos estimadores. As principais diferenças entre a regressão Ridge e o LASSO estão na penalização utilizada na expressão da soma de quadrados dos erros que deve ser minimizada para estimar os coeficientes de regressão. Portanto, a regressão Ridge levar as estimativas de alguns coeficientes à zero, enquanto, que o LASSO faz com que as estimativas de alguns coeficientes sejam exatamente iguais a zero, tornando o segundo modelo mais interpretável, com menor número de co-variáveis.

O objetivo deste projeto é propor modelos que combinam extensões do ELM com modelos de regressão: Ridge e LASSO, para caracterizar o problema a partir de uma visão analítica, integrando e estruturando, entre outras, as bases de dados: DataSUS e IBGE que englobam aspectos desde os sanitários até os econômico-sociais da população do agreste do estado de Pernambuco [4].

Este projeto de mestrado objetiva investigar e propor a combinação de técnicas de ELM com modelos de regressão aplicados ao problema diagnóstico de doenças, como por exemplo, decisões de tratamento, assistência ao paciente, decisões financeiras e operacionais. Espera-se que a aplicação da técnica de mineração de dados permita a criação de um modelo de previsão utilizando combinação ELM com regressores, a fim de promover avanços na área de saúde. Durante o estudo, um ambiente experimental para avaliação do modelo proposto será implementado, como também, serão utilizados dados (artificiais/reais) e o desempenho do modelo será medido pelo erro de predição estimado pelo método Monte Carlo no conjunto de teste utilizando a Linguagem Python.

### Referências Bibliográficas

- [1] Y. Bengio. Learning Deep Architectures for AI. in Foundations and Trends in Machine Learning, 2(1):1–127, 2009.
- [2] G. Huang, G. B. Huang, S. Song, K. You. Trends in Extreme Learning Machines: A Review. Neural Networks, 61, 2015, pp 32-48.
- [3] MONTGOMERY D.C.; PECK, E.A. e VINING, G. G. Introduction to Linear Regression Analysis, Wiley-Interscience, 2006.
- [4] Dados Caruaru. Disponível em <https://caruaru.pe.gov.br/portal-da-transparencia/dados-abertos/>. Acesso em 21/06/2021.