

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO

Pró-Reitoria de Pesquisa e Pós-Graduação

Departamento de Pesquisa

DESCRIÇÃO PARCIAL DAS METAS ALCANÇADAS

IDENTIFICAÇÃO DO PROJETO DE PESQUISA

Grande Área CNPq: Ciências Exatas e da Terra

Área CNPq: Probabilidade e Estatística

Ciência de Dados e Aprendizado Estatístico Aplicados à Saúde

Nome do grupo de pesquisa: DasLab - Laboratório de Ciência de Dados

Linha de pesquisa: Ciência de Dados e Aprendizado Estatístico Aplicados à Saúde

Pesquisador responsável (orientador): Prof^a. Dr^a. Agatha Sacramento Rodrigues

Interpretabilidade em modelos preditivos na área da saúde

Nome do aluno: Ornella Scardua Ferreira

Curso e período: Estatística / 5º

Introdução e justificativa:

Em *Machine Learning* (ML, em português Aprendizado de Máquina), as análises são embasadas por modelos estatísticos e algoritmos computacionais, muita vezes, sofisticados, cujos resultados, predições ou decisões, decorrem do aprendizado direto sobre os dados. No entanto, alguns modelos que têm ótimo desempenho preditivo e acurácia podem não ser autoexplicáveis. Por essa razão, a utilização de medidas de interpretabilidade se faz necessária se há a pretensão em entender as decisões que foram tomadas pelos algoritmos.

Nos dias de hoje, modelos de ML têm sido usados para amparar serviços considerados essenciais para a sobrevivência humana e, portanto, se tornado bastante populares nas diversas áreas de conhecimento, incluindo a da saúde. Em particular, na medicina clínica, o desejo em prever e depreender desfechos de interesse tem sido cada vez mais impulsionado com o advento da área de *Machine Learning* e técnicas de interpretação.

Nesse sentido, este projeto visa discutir modelos preditivos e métodos de interpretabilidade na área da saúde como resultado de uma aplicação real na área da medicina obstétrica, em que se espera prever se uma gestante diagnosticada com diabetes gestacional fará o uso de insulina em algum momento precedente ao parto e explicar o motivo por que ocorreu tal desfecho.

Materiais e métodos utilizados:

A etapa primeira consistiu no processamento e análise de consistência e exploratória dos dados e a segunda, na modelagem. A amostra original foi dividida em 75% dos dados para treinar e os 25% restantes para testar. Na amostra de teste é avaliada a acurácia do modelo, percentual de acertos conforme um limiar (James & Tibshirani, 2013). Na amostra de treinamento os melhores parâmetros, cujos valores são determinados pelo menor erro, são encontrados via método de validação *k-fold* e usados no ajuste na amostra de treinamento (Izbicki, 2020).

Com base nas literaturas de Morettin (2020) e Izbicki (2020), os modelos considerados aqui são apresentados na primeira coluna (“modelo”) da Tabela 1, todos ajustados por meio de pacotes do R, versão 4.0.3, linguagem computacional usada neste projeto, sob a IDE RStudio. No modelo de regressão logística a curva ROC (*Receiver Operating Characteristic*) foi usada para determinar o limiar de classificação que melhor otimizava as medidas de sensibilidade e o complementar da especificidade.

Já Amorim (2019) e Molnar (2019) foram quem fundamentaram a parte teórica e prática acerca dos métodos de interpretabilidade globais e individuais, como o Gráfico de Dependência Parcial e *SHapley Additive exPlanations*.

Resultados e discussão:

Inicialmente, foi realizada uma breve análise exploratória em cima de uma base de dados composta por 31 variáveis e 408 observações, cujos indivíduos são gestantes diagnosticadas com diabetes gestacional que realizaram o pré-natal no Hospital das Clínicas da USP. As variáveis preditoras selecionadas para o processo seguinte de modelagem foram 9: idade, número de gestações, IMC, histórico de diabetes na família, antecedente de macrosomia, antecedente de diabetes gestacional, indicador de tabagista, indicador de hipertensão e valor da glicemia de jejum. Dentre as quais, o número de observações passou a ser 404 após a exclusão de dados faltantes.

Como resultado, a acurácia conjuntamente à função e ao pacote associados a cada um dos modelos preditivos considerados estão na Tabela 1 que segue.

modelo	acurácia	função	pacote
Regressão Logística	0,792	glm	R Base
Análise Discriminante Linear	0,712	lda	MASS
Análise Discriminante Quadrática	0,722	qda	MASS
k Vizinhos Mais Próximos	0,613	knn	class
Support Vector Machines Linear	0,653	svm	e1071
Support Vector Machines Gaussiano	0,683	svm	e1071
Support Vector Machines Polinomial	0,732	svm	e1071
Árvores de Classificação	0,693	rpart	rpart
Bagging	0,623	train	caret
Florestas Aleatórias	0,663	train	caret
XGBoost	0,795	train	caret

Tabela 1: Acurácia e funções e pacotes com os quais os modelos foram ajustados.

Tendo em vista a maior acurácia, o modelo previamente ajustado escolhido para serem feitas as aplicações das métricas de interpretabilidade é o *XGBoost*, em que os resultados iniciais podem ser encontrados em http://rpubs.com/orncar/interpret_methods_xgb. Todas as técnicas de interpretação foram empregadas via pacote *iml* (veja sua documentação aqui).

Metas futuras:

Nos próximos 6 meses de pesquisa é esperado

1. que os métodos de interpretabilidade estudados sejam documentados, discutindo suas vantagens e desvantagens;
2. que seja desenvolvida uma plataforma web por meio de um aplicativo Shiny (<https://www.shinyapps.io/>)

- para melhor acessibilidade dos usuários aos métodos de interpretabilidade;
3. que sejam feitas reuniões com os pesquisadores dos ambulatórios de diabetes gestacional do Departamento de Obstetrícia da FM-USP para, num primeiro momento, discutir os dados e, num segundo momento, apresentar os resultados; e
 4. que o relatório final da pesquisa seja finalizado e entregue.

Referências

- Amorim, W. N. d. (2019). *Ciência de dados, poluição do ar e saúde*. PhD thesis, Universidade de São Paulo.
- Izbicki, Rafael.; Santos, T. M. d. (2020). *Aprendizado de máquina: uma abordagem estatística*. Câmara Brasileira do Livro, SP, Brasil, 2ª edition.
- James, Gareth; Witten, D. H. T. & Tibshirani, R. (2013). *An Introduction to Statistical Learning with Applications in R*. Springer.
- Molnar, C. (2019). *Interpretable Machine Learning - A Guide for Making Black Box Models Explainable*.
- Morettin, Pedro A.; Singer, J. M. (2020). *Introdução à Ciência de Dados - Fundamentos e Aplicações*.