Universidade Estadual de Campinas

Ciência e Visualização de Dados em Saúde

Prof. Dr. André Santanchè, Profª. Drª. Paula Dornhofer P. Costa
Profª. Drª. Leticia Rittner e Profª. Drª. Taís Freire Galvão

Falhas de Diagnóstico Médico e os Impactos no Sistema Único de Saúde Brasileiro

Alunos: Gleyson Roberto do Nascimento. RA: 043801.
Negli René Gallardo Alvarado. RA: 234066.
Rafael Vinícius da Silveira. RA: 137382.
Sérgio Sevileanu. RA: 941095.

Junho/2021

O Problema

As falhas de diagnóstico médico são responsáveis, num primeiro momento, por consequências aos indivíduos que foram vitimadas por estas falhas, todavia, até onde se estendem essas consequências?

Neste contexto, o projeto busca compreender melhor quais são os impactos que as falhas trazem ao Sistema Único de Saúde, se é possível mitigar os resultados negativos e auxiliar na predição de casos de falha de diagnóstico através do Aprendizado de Máquina.

As Perguntas de Projeto

1) Qual é a extensão do impacto causado pelas falhas de dignóstico médico dentro do Sistema Único de Saúde do Brasil?

2) Uma vez conhecida esta extensão, é possível propor um método de predição de casos de falha de diagnóstico em Machine/Deep Learning?

3) Sendo possível essa predição, esse método de fato seria útil, de fácil acesso e amigável para os profissionais do SUS?

O Método e os Dados

Através da metodologia CRISP-DM, os dados coletados serão reunidos, analisados, limpos, minerados (nestas etapas, haverá a ênfase na análise estatística) e, após isso, através de aprendizado de máquina, se for possível, será criado um modelo de predição de falhas de diagnóstico médico e por fim este modelo será analisado e, se for validado, pode ser implementado como auxílio para profissionais de saúde do SUS.

Para este fim, foram elencadas as bases de dados do Sistema de Informações Hospitalares do SUS e o CID - 10.

Dificuldades Enfrentadas

Durante o processo, a maior dificuldade enfrentada pelo grupo foi lidar com o Big Data envolvendo o banco de dados do Sistema de Informações Hospitalares (SIHSUS) e com os recursos computacionais disponíveis.

De forma geral, para lidar com essa dificuldade, o grupo utilizou artifícios como a computação paralela, através da biblioteca Dask, a mudança do Kernel do Google Colab para uso mais intenso da GPU através do RAPIDS e dataframes com uso do CUDA.

A segunda grande dificuldade foi encontrar um bom modelo de ML que trouxesse não só uma boa acurácia, mas também sensibilidade e especificidade aceitáveis para representar a classificação dos diagnósticos e evitar *overfitting* e *underfitting*.

Mudanças de Percurso

Durante a análise exploratória, o grupo ainda não tinha experiência em lidar com o Big Data, desta forma, algumas mudanças foram necessárias:

1) Ao invés de analisar o SUS como um todo, analisar os maiores Estados de cada Região do Brasil;

2) Descartar bases secundárias que poderiam melhorar as features e relações estatísticas, mas que custariam em termos de recursos computacionais;

3) Usar metodologias de ML compatíveis com Dask e CUDA.

Referenciais Teóricos

Foram utilizados os seguintes referenciais teóricos:

[1] Martinez, F.; Contreras, L. "**CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories**". IEEE Transactions on Knowledge and Data Engineering, 2020;

[2] Singh, H; Onakpoya, I. **"The global burden of diagnostic errors in primary care"**. BMJ Qual Saf, 2017.

[3] Daniel, J. "**Data Science with Python and Dask**". 1st Edition, Manning Publisher, 2019;

[4] **https://rapids.ai/index.html** <acesso em 01/06/2021>.

Ferramentas Utilizadas

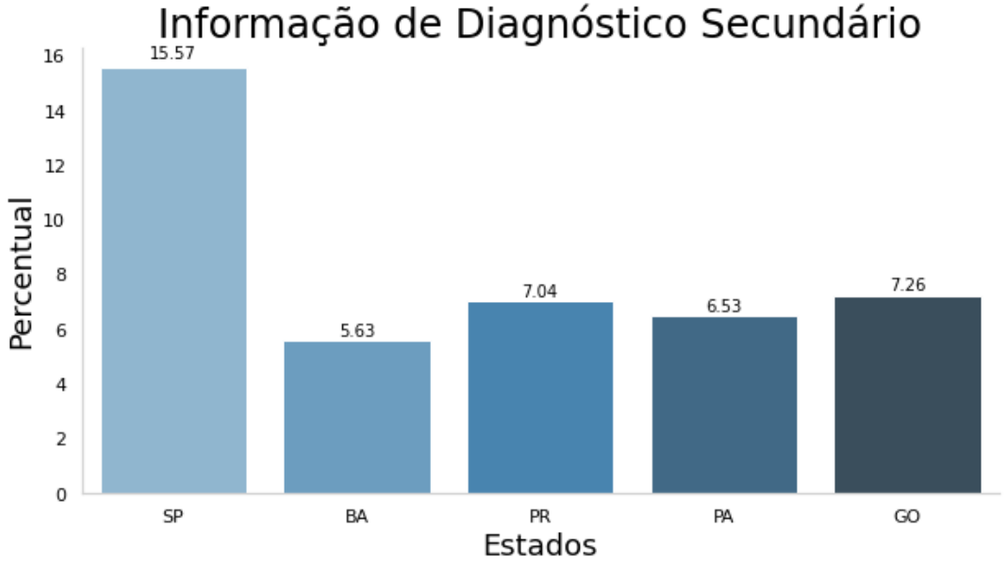
Para este projeto as principais ferramentas utilizadas foram:

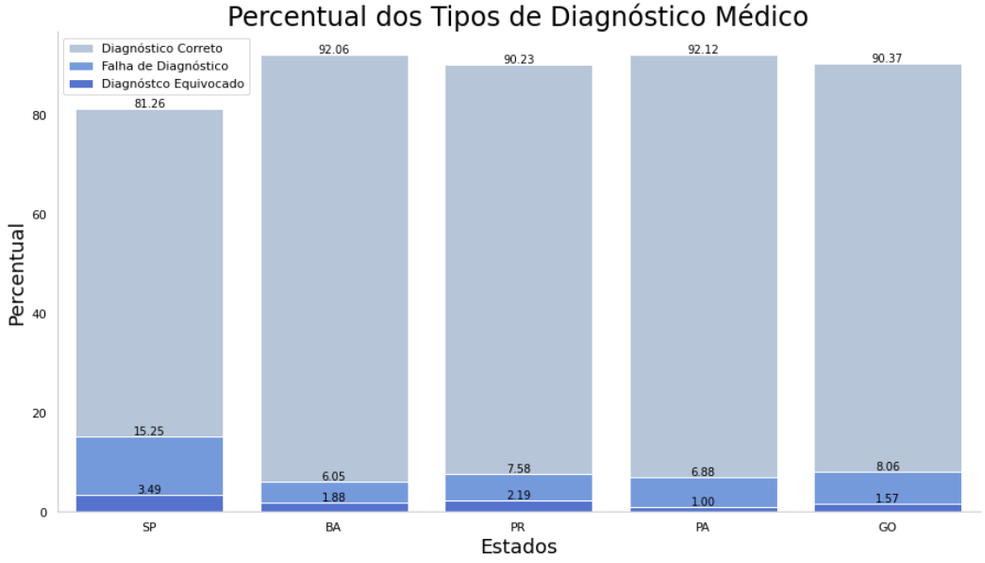
A) **Rapids**: No Google Colab ele altera o kernel, fazendo com que os processos sejam otimizados com a utilização forçada da GPU através do CUDA. Em https://rapids.ai/start.html existe um notebook Colab com as instruções de instalação;

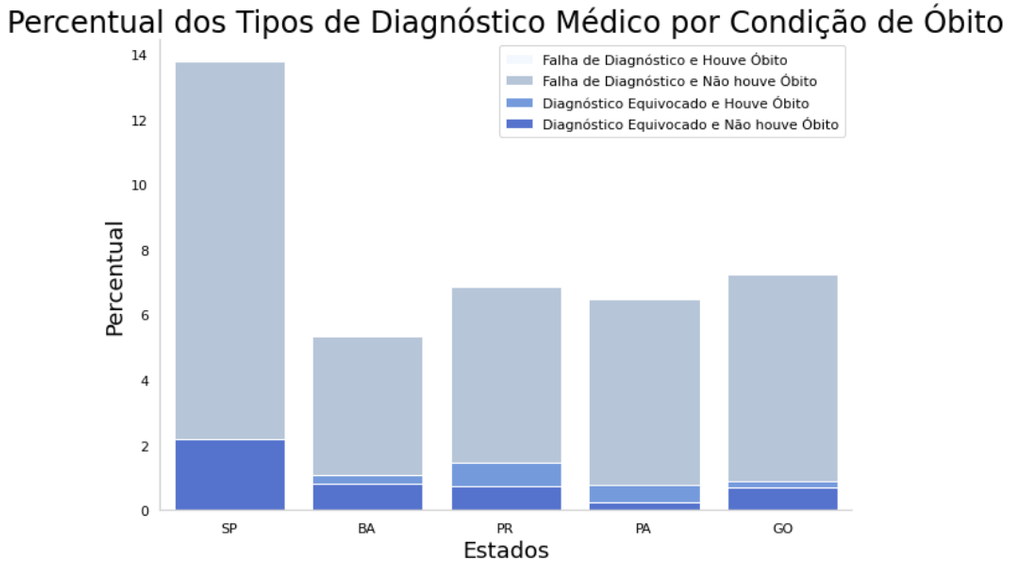
B) **Dask**: Esta biblioteca implementa computação paralela em Python, de forma que para Big Data é fundamental, uma vez que os recursos computacionais disponíveis não são suficientes para lidar com estes arquivos. Além disso, substitui perfeitamente o Pandas através do Dask Dataframe e possui um módulo de Machine Learning específico. Em dask.org é possível ver toda a documentação da biblioteca;

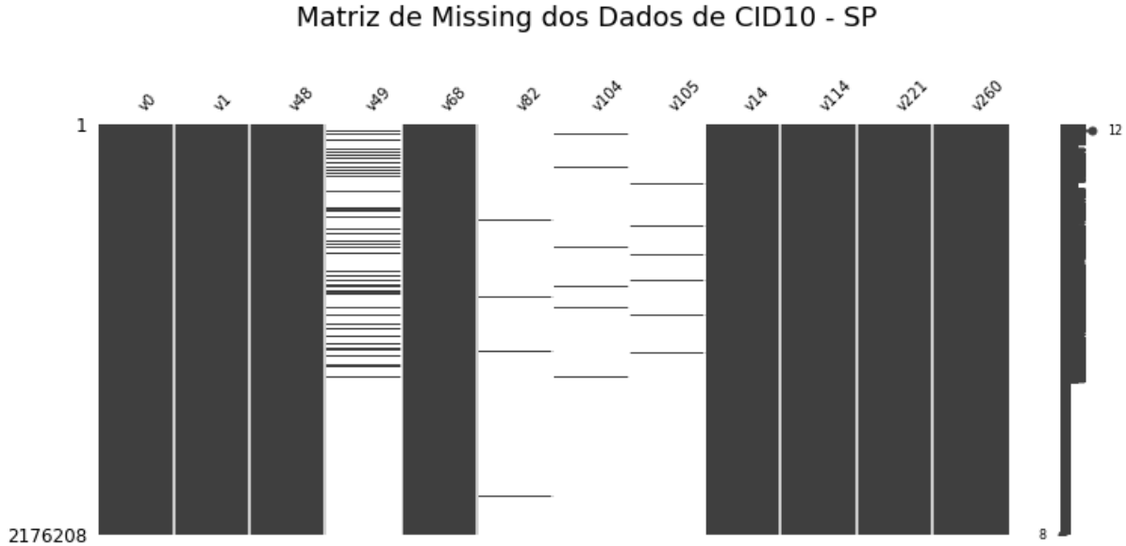
C) **Scikit-Learning**: Biblioteca clássica para Machine Learning, os modelos de aprendizagem se adequaram perfeitamente ao uso em computação paralela. Através do scikit-learn.org é possível ver toda a documentação da biblioteca.

Resultados Estatísticos

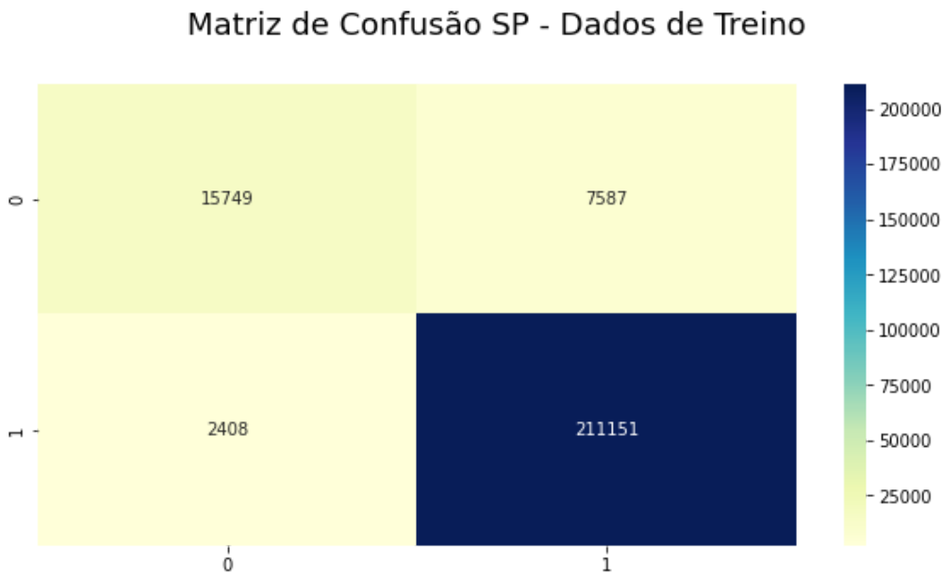


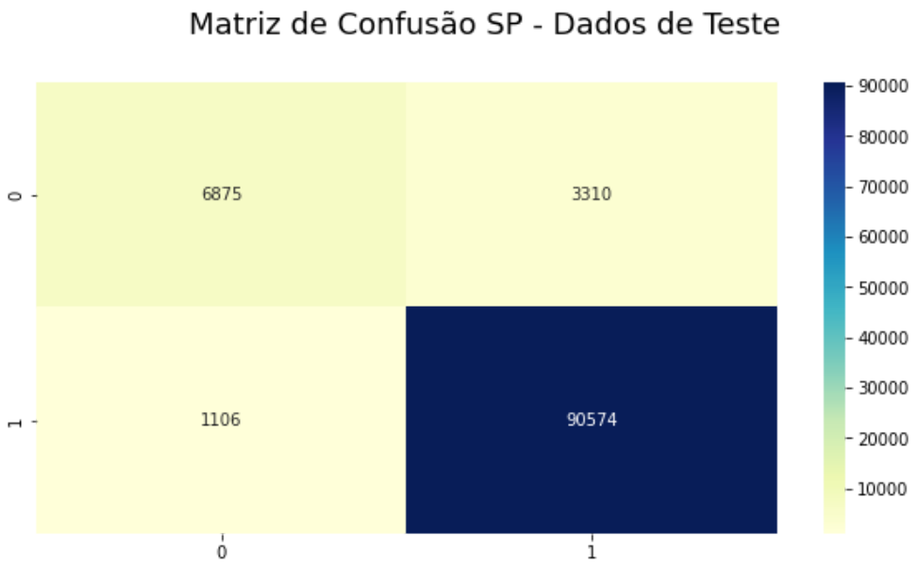


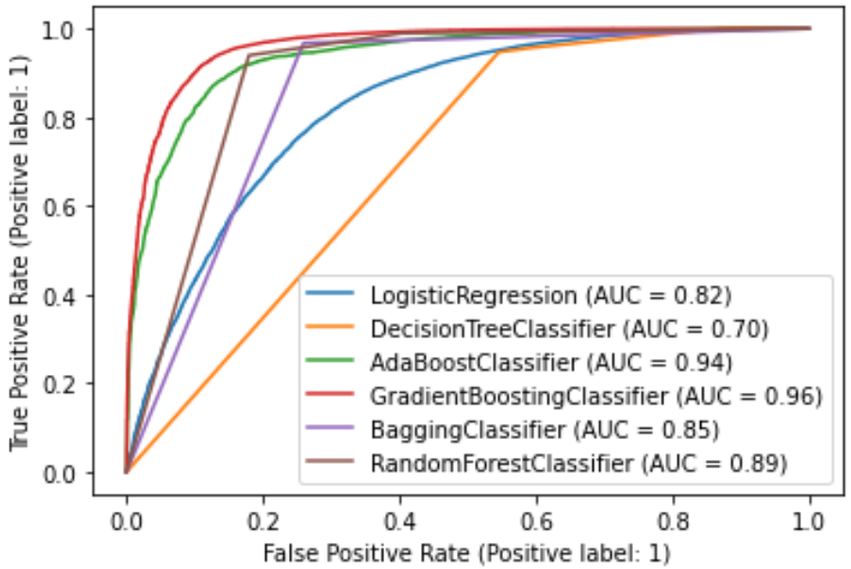




Resultados Aprendizagem de Máquina I







Parâmetro ▲	Treino	Teste
Acurácia	0.957808311699276	0.95664850537476
Especificidade	0.99	0.99
RMSE	0.0421916883007239	0.0433514946252392
Sensibilidade	0.67	0.68

Resultados Aprendizagem de Máquina II

Estado	Acurácia	Sensibilidade	Especificidade	RMSE
SP - treino	0.957808311699276	0.67	0.99	0.0421916883007239
SP - teste	0.95664850537476	0.68	0.99	0.0433514946252392
PR - treino	0.992274580233187	0.79	1	0.00772541976681205
PR - teste	0.983747357293869	0.61	0.99	0.016252642706131
PA - treino	0.999900616179686	0.99	1	0.0000993838203140528
PA - teste	0.99478563151796	0.7	1	0.00521436848203939
GO - treino	0.989533043936873	0.91	0.99	0.0104669560631262
GO - teste	0.981305803571428	0.84	0.99	0.0186941964285714
BA - treino	0.988455428067078	0.84	1	0.0115445719329214
BA - teste	0.980856586632057	0.74	0.99	0.0191434133679428

Importância	Variável	Descrição
1	v69	Definição de óbito
2	v51	Definição do motivo de saída/permanência
3	v48	Código do diagnóstico principal (CID10)
4	v5	Número da AIH
5	v100	Código CNES do hospital
6	v7	Definição do número da AIH
7	v117	Sequencial da AIH na remessa
8	v68	Indica óbito
9	v79	Definição do número de filhos do paciente
10	v249	Dias de permanência

Discussão de Resultados I

A) **Estatísticos**: Os principais resultados demonstram que as falhas de diagnóstico e os diagnósticos equivocados realmente são exceções e ocorrem mais para homens brancos e pardos (quando há declaração de raça), fato decorrente de fatores sócio-econômicos, uma vez que esse grupo tem maior acesso aos hospitais; o tempo médio de internação geralmente dobra quando há falha de diagnóstico e esta leva ao óbito do(a) paciente; em termos de óbito, as falhas de diagnóstico tem menor representatividade que os diagnósticos equivocados; no geral, neoplasias são as doenças com maior incidência de falhas de diagnóstico. Por fim, no banco de dados há falta de dados de diagnósticos secundários a partir de 2015;

Discussão de Resultados II

B) ***Machine Learning***: Através do modelo de *Gradient Boosting* foi possível criar um classificador para analisar a importância das *features* que levam a uma falha de diagnóstico ou diagnóstico equivocado, contudo, devido ao enviesamento ocorrido quando a classe "diagnóstico correto" estava presente, não foi possível elaborar um classificador funcional para predizer as 3 classes de diagnóstico. Com relação as *features*, basicamente existem 3 categorias importantes: a primeira tem relação com o hospital, alguns possuem maiores índices de falha/equívoco que outros; a segunda com as condições técnicas (diagnóstico principal, tempo de internação, óbito) e a terceira com as condições sócio-econômicas do paciente (quantidade de filhos);

Conclusão

Neste projeto, foi possível concluir que:

1) De forma geral, as falhas e equívocos de diagnóstico são sim exceções, contudo, há uma grande divergência com relação ao real valor, uma vez que o banco de dados do SIHSUS é mantido e modificado por muitos usuários e com diferentes realidades, sujeito a erros sistemáticos, além disso, não há dados de diagnósticos secundários a partir de 2015 em todo o banco, o que prejudica a transparência do processo;

2) Em termos humanos, os diagnósticos equivocados levam mais ao óbito, fato que não era exatamente esperado, já que o diagnóstico dado era próximo do real, assim, uma investigação mais detalhada no hospital é justificada;

3) Em termos econômicos, as falhas de diagnóstico geralmente dobram o tempo de internação e, além disso, possuem maiores custos médios por internação, causando maior impacto ao sistema.

Trabalhos Futuros

São possibilidades de trabalhos futuros:

1) Análise dos diagnósticos médicos através do uso do *Deep Learning* para *Big Data*: desta forma talvez fosse viável um classificador funcional para as 3 classes de diagnóstico e que poderia ser implementado no SUS para o auxílio de diagnóstico;

2) O SIHSUS é uma base que representa apenas os hospitais, de forma que há uma gama muito grande de unidades de atendimento não contempladas, assim, se houver a possibilidade da inclusão de mais databases com outros tipos de unidades, o modelo teria um retrato mais fiel do sistema;

3) Ter acesso à algumas *features* como exames pedidos e resultados obtidos poderiam tornar o modelo mais completo e confiável.