**Título**:

O uso de sistemas de Inteligência Artificial para predição de Doença Renal Crônica: um desenvolvimento de modelos computacionais com dados de pacientes acompanhados pelo Estudo Longitudinal de Saúde do Adulto – ELSA-Brasil.

**Introdução (referencial teórico e justificativa):**

Ultimamente, com o envelhecimento populacional e as crescentes taxas de doenças crônicas não transmissíveis, tem-se posto em pauta a importância do desenvolvimento de métodos seguros, eficazes, econômicos e que descentralize essas análises de grandes centros laboratoriais, ou seja, interiorize e cubra uma maior parcela da população (BIKBOV et al., 2020). Nesse sentido, uma dessas pautas é a Doença Renal Crônica (DRC), afecção nos rins, caracterizada pela perda progressiva e irreversível da massa nefrônica responsável pela filtração sanguínea e excreção de metabólitos do organismo, que tem direta relação com o envelhecimento, bem como, com comorbidades ligadas a ele, a exemplo do diabetes mellitos e a hipertensão arterial sistêmica (KDIGO, 2012).

Nos últimos anos, muitos métodos têm sido desenvolvidos para o monitoramento renal, em especial para a averiguação da massa nefrônica funcional e monitoramento da Taxa de Filtração Glomerular (TFG) parâmetro mais utilizado para esse fim. Dentre esses métodos, os mais interessantes, tendo em vista sua praticidade e custo em relação aos exames de imagem e de excreção de marcadores como a inulina e iotalamato, são os cálculos que estimam a TGF a partir de dados sociodemográficos (idade, sexo, cor da pele) e valores de creatinina urinária ou clearence de creatinina. Entretanto, foi questionada a precisão das equações CKD-Epi (LEVEY et al, 2009) e MDRD-4 (LEVEY et al, 2006), que atualmente são as formas mais disseminadas de monitoramento renal no que diz respeito a detecção precoce, mais especificamente a confiança de que os pesos presentes nas equações irão ter o poder de estimar a TGF de todas as pessoas com a mesma exatidão, logo, se não vai superestimar ou subestimar e valor.

Assim, os modelos computacionais apresentam-se como uma alternativa para sanar essa lacuna, visto que, modelos de Redes Neurais Artificiais (RNA) e Aprendizagem de Máquina ou Machine Learning (MA), são alguns sistemas de Inteligência Artificial (IA) (KOVÁCS, 2002), que treinam um modelo matemático a partir de um resultado preestabelecido (como um comparativo com padrão ouro), os pesos desses modelos são ajustados entao a cada ciclo de testes que o modelo é submetido e ao final de cada ciclo os valores estimados são submetidos a equações que averiguam o erro das predições. Algoritmos criados nessa arquitetura, tem inúmeras aplicações para predições e análise de padrões, são vastamente utilizados para análise de padrões, por exemplo em imagens de câmera de segurança ou digitalização de símbolos, números, alfabetos e etc, bem como há aplicações na área da saúde também, como descrito no trabalho de Santos et al, 2005, que avaliou o uso desse sistema para analisar a soroprevalência de hepatite A em Duque de Caxias, no estado do Rio de Janeiro.

**Hipótese da pesquisa e principais perguntas**:

1. Por causa do ajuste de pesos e do sistema de correções de erros, o algoritmo desenvolvido nessa pesquisa estimará melhor os valores de TFG que as equações vigentes?
2. Sendo uma ferramenta computacional, será possível fazer mais implementações no script de programação?
3. De uma forma diferencial dessa pesquisa em relação às equações, é possível adicionar mais filtros que enriqueçam o poder de predição do algoritmo, como por exemplo, diferenciar grávidas de pessoas obesas, neuropatas, hepatopatas, hipertensos, diabéticos de indivíduos normais.
4. Como essa é uma forma inédita de predição nesse campo, seria possível o sistema traçar novos padrões e relações que não foram descritas antes, como o impacto de hábitos alimentícios, nocivos (drogas licitas ou ilícitas), cotidianos (estado de sono, sedentarismo, exercícios físicos, dentre outros) na saúde renal desse paciente?

**Metodologia (plano de análise, métodos estatísticos):**

Utilizaremos as ferramentas computacionais Excel, Python e Anaconda Navigator para a limpeza e preparação dos dados, logo em seguida será escrito o script para a submissão do banco de dados aos cálculos da RNA, acoplada ao sistema de correção de erros que reajustará os pesos da equação a cada ciclo. Quando esses primeiros treinamentos forem finalizados, submeteremos uma nova parcela da amostra que irão testar a acurácia do modelo, os dados gerados por cada ciclo de treinamento e testagem serão coletados para a análise comparativa com as equações e o padrão ouro, para posterior discussão. Programas como o SPSS, Graphpad, STATA, ANOVA, R Studio, poderão ser utilizados para algumas análises estatísticas (Correlações, Regressão Linear, comparação de métodos por Bland-Altman) do método e geração de gráficos que também serão materiais de discussão. A inserção e

**Metodologia (variáveis utilizadas)**

* Variáveis referentes aos bancos de dados de urinas validadas do ELSA-Brasil.
* Variáveis antropométricas, hemodinâmicas, bioquímicas, presença de comorbidades e estilo de vida (atividade física e tabagismo) dos indivíduos que participam das pesquisas relacionadas com analise de função renal, ou mesmo que envolvam parâmetros séricos e urinários.

**Metodologia (população toda ou amostra):**

População toda que foi submetida a análise de urina do ELSA-Brasil.

**Material biológico**

No  
  
**Comentários adicionais:**

No

**Referências**  
KOVÁCS, Zsolt László. **Redes neurais artificiais**. Editora Livraria da Fisica, 2002. Disponível em: ([Link](https://books.google.com.br/books?hl=pt-BR&lr=&id=O0nLxR67wmUC&oi=fnd&pg=PA9&dq=redes+neurais+artificiais&ots=hd_CUH3Pjg&sig=QPS3OKXCkSPH-D8WmSmmiB_gxfU#v=onepage&q=redes%20neurais%20artificiais&f=false)). Acesso em 18 de agosto de 2021.

BIKBOV, B. et al. **Global, Regional, and National Burden of Chronic Kidney Disease, 1990–2017: A Systematic Analysis for the Global Burden of Disease Study 2017.** The Lancet, v. 395, n. 10225, p. 709–733, 29 fev. 2020.

KIDNEY DISEASE: IMPROVING GLOBAL OUTCOMES (KDIGO). KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. v. 3, p. 1–150, 2013.

LEVEY, A. S. et al. A New Equation to Estimate Glomerular Filtration Rate. **Annals of Internal Medicine**, v. 150, n. 9, p. 604–612, 5 Maio 2009.

\_\_\_\_\_\_. Using Standardized Serum Creatinine Values in the Modification of Diet in Renal Disease Study Equation for Estimating Glomerular Filtration Rate. **Annals of Internal Medicine**, v. 145, n. 4, p. 247–254, 15 ago. 2006.

SANTOS, A. M. et al. **Usando redes neurais artificiais e regressão logística na predição da Hepatite A. Revista Brasileira de Epidemiologia**. 2005, v. 8, n. 2, pp. 117-126. Epub 12 Jun 2007. ISSN 1980-5497. Disponível em: ([Link](https://doi.org/10.1590/S1415-790X2005000200004)). Acesso em 18 de agosto de 2021.

**AUTORES**André Paulino de Lima - [andre.p.lima@usp.br](mailto:andre.p.lima@usp.br) (USP)

Brunela Della Maggiori Orlandi - [brunella@alumni.usp.br](mailto:brunella@alumni.usp.br) (IFNMG)

Lísias Pereira Novo - [lisias.novo@ufob.edu.br](mailto:lisias.novo@ufob.edu.br) (UFOB)

Sérgio Santos Evangelista – [sergio.e2981@ufob.edu.br](mailto:sergio.e2981@ufob.edu.br) (UFOB)

Wagner Luis da Cruz Almeida - [walmeida@ufba.br](mailto:walmeida@ufba.br) (UFBA)

**ORIENTADOR**José Geraldo Mill - josegmill@gmail.com (UFES)