

Predição de Prognóstico Clínico

Bruna Almeida Osti* e Fabio Fogliarini Brolesi†

Instituto de Computação, Universidade Estadual de Campinas - UNICAMP
Campinas, SP

Email: *b231024@dac.unicamp.br, †brolesi@gmail.com

Resumo—Este trabalho busca abordar o planejamento do projeto da disciplina, explanando a especificação do problema, estruturação da metodologia, questões éticas, como será feito a avaliação da abordagem e outros detalhes relevantes.

Index Terms—Trustworthy, AI, Ethical

I. INTRODUÇÃO

O efeito do avanço tecnológico fez possível a criação de ferramentas para coleta e processamento de dados nunca antes vistos, que são usados para dar vida a sistemas automatizados e auxiliar a tomada de decisões de diversos problemas.

Em modelos de machine learning os algoritmos baseiam seu aprendizado diretamente aos dados que são apresentados a ele que especificam quais são os outputs corretos. Entretanto, por replicar e generalizar dados de grupos majoritários acabamos replicando vieses sociais que já pré-existem na nossa sociedade [1]. Um exemplo muito conhecido é a ferramenta de recrutamento da Amazon que continha viés de gênero e foi descontinuada [2].

A confiabilidade, segurança, proteção e transparência necessárias para confiarmos nesses algoritmos em todos os casos permanecem indefinidas. Como resultado, há uma crescente vontade popular de ter mais justiça, robustez, interpretabilidade e transparência nesses sistemas [3], consequentemente aumentar a confiabilidade.

A Inteligência artificial confiável (*Trustworthy AI*) sugere um sistema que é legal, eticamente aderente e tecnicamente robusta. E que garante todo seu potencial estabelecendo a confiança em todos os estágios do ciclo de vida, desde o design até o uso. Por outro lado, existem alguns desafios acerca da implementação e da garantia que esses requisitos sejam devidamente cumpridos, deve haver a decisão de ser vulnerável levando em conta que o sistema pode ter vieses e baseando-se em sua confiabilidade e em outros fatores [3].

Afim de evitar os riscos potenciais da implantação de sistemas com IA, a Lei de Proteção de Dados (LGPD) foi sancionada recentemente para regular a proteção de dados no país, que tem como principal motivação proteger os direitos fundamentais como liberdade, privacidade, honra, imagem, entre outros, voltado para essas necessidades atuais como à prevenção e correção de vieses algorítmicos e seus impactos negativos na vida das pessoas [4].

II. ESPECIFICAÇÃO DO PROBLEMA

O problema é identificar quais pacientes terão como prognóstico alta de unidade de terapia intensiva em até 7 dias após a cirurgia baseado em características presentes no conjunto de

dados. Os prognósticos em contextos clínicos tem por objetivo ajudar clínicos, pacientes e familiares a tomarem decisões esclarecidas a respeito de cuidados de saúde com base em informações disponíveis sobre cada paciente no presente para prever desfechos no futuro, conforme [5].

III. TECNOLOGIAS

Para o presente problema, pretende-se estruturar uma automação que seja capaz de prever quais pacientes sairão da unidade de terapia intensiva dadas características presentes nos datasets fornecidos que tem relação com cirurgias ortopédicas e dados de estadia em unidades de terapia intensiva, todos eles anonimizados. Os modelos prognósticos usam vários fatores em combinação para prever o risco de resultados clínicos futuros em pacientes. Um bom modelo deve (i) fornecer previsões precisas que informam os pacientes e seus cuidadores, (ii) apoiar a pesquisa clínica e (iii) permitir decisões para melhorar os resultados dos tratamentos aos pacientes, segundo [5]. Um modelo prognóstico tem três fases principais: desenvolvimento do modelo (incluindo validação interna), validação externa e investigações de impacto na prática clínica. Embora muitos modelos prognósticos sejam propostos, poucos são atualmente usados na prática, conforme mostra [5].

Entre os prognósticos existentes, alguns são apresentados abaixo:

- **Simplified Acute Physiology Score III (SAPS3):** Considera a idade, se já vêm de pré-hospitalização, qual a localização anterior (urgência, UCI e outros), se possui alguma comorbidade, qual o motivo da entrada, e variáveis medidas no momento do atendimento, para gerar a mortalidade prevista [6].
- **Palliative Prognostic Index (PPI):** Prediz a mortalidade em pacientes terminais com base em cinco critérios, sendo eles a escala de desempenho paliativo, a ingestão oral, presença de edemas, se há dispneia em repouso ou se há delírios. Com base nestes parâmetros, é informado em até quantas semanas é esperado que este paciente venha à óbito [7].
- **Palliative Prognostic Score (PaP):** Semelhante ao anterior, também foca em pacientes em cuidados paliativos. Neste caso os parâmetros de entrada são Dispneia, Anorexia, Karnofsky Performance Status, Previsão Clínica de Sobrevivência (semanas), Total de WBC e Linfócitos %, para prever a probabilidade de sobrevivência em 30 dias [8].

- **MAGGIC Risk Calculator for Heart Failure (MAGGIC RCHF):** Este prognóstico foca apenas em pessoas que sofrem da condição de insuficiência cardíaca, e com base nos parâmetros de entrada, estima a mortalidade em 1 e 3 anos [9].

Este tipo de problema encontra-se na categoria de ciências da decisão, no campo da medicina, mas também é objeto de estudo do campo da administração hospitalar e também de custos para o sistema de saúde uma vez que a classificação do prognóstico pode apoiar em tomadas de decisões reduzindo custos ou trazendo luz a casos específicos que eventualmente podem requerer maior atenção. Segundo [10] ter um prognóstico bem definido ajuda os pacientes e os familiares a tomar decisões adequadas a respeito do fim da vida daqueles cujo risco de morte é muito alto e a identificar intervenções personalizadas para evitar futuras hospitalizações.

O resultado da avaliação do modelo será satisfatório na medida em que ele for capaz de apoiar a decisão médica para que ela seja tomada no menor tempo possível e com a melhor resposta possível para o paciente afim de que os cuidados necessários para aquele paciente sejam ministrados o quanto antes (seja para alguma alta, sejam cuidados paliativos) afim de trazer conforto para o paciente.

A Inteligência artificial confiável (*Trustworthy AI*) não implica automaticamente que o sistema é confiável, sugere uma relação entre um fiduciante (trustor) e um fiduciário (trustee), visto que existem certas propriedades que o tornam confiável, entretanto, não implica que é confiável. O trustor deve tomar a decisão de ser vulnerável levando em conta vieses que trustee possa ter, baseado em sua confiabilidade e em outros fatores [3].

Conforme [3] é possível listar vários atributos de pessoas e sistemas de informação confiável e segundo [11], utilizaremos os seguintes atributos:

- **habilidade:** conforme [11], é "grupo de habilidades, competências e características que permitem que uma parte tenha influência dentro algum domínio específico". Neste sentido, entendemos que há um conjunto de habilidades dentro do contexto da prática médica que deve, em algum grau ser levado em consideração no momento da análise do modelo.
- **integridade:** segundo [11], é preciso uma aderência a "um conjunto de princípios que o credor considera aceitável". Isso requer confiança de que o agente fiduciário agirá de acordo com um conjunto de princípios e que estes se alinham com os valores do credor. No caso do presente trabalho, a proposta é manter uma estrutura íntegra e que garanta integridade e entendimento de todas as partes interessadas no processo, afim de garantir esta integridade.
- **benevolência:** [11] define benevolência como "a medida em que se acredita que um curador quer fazer o bem ao administrador"
- **previsibilidade:** [11] diz que a previsibilidade reforçará as percepções dos três outros aspectos do fiduciário (trustee). No caso da previsibilidade, é importante que

a confiança seja mantida ao longo do tempo por meio de relacionamentos.

IV. ÉTICA

Do ponto de vista ético, especificamente pensando numa ética consequencialista, podemos pensar em resultado de prognóstico, sem fazer qualquer suposição acerca do paciente, apenas utilizando os métodos de predição propriamente ditos, mas que claramente tem uma acurácia percentual abaixo de 100%. Segundo [12], "Uma teoria de tipo consequencialista deve estar comprometida, pelo menos em alguma medida, com o ideal do controle das variáveis que maximizam o bem, seja ele definido em termos de utilidade social, seja em termos da excelência de uma elite de indivíduos". Um exemplo seria o caso de, a partir de característica do paciente, e com resultados de prognósticos de evolução para óbito independente de uma cirurgia ortopédica, a decisão de não realizar a operação.

Por outro lado, entendendo a possibilidade de uso de uma ética utilitarista, que é derivada da mesma ética consequencialista, na medida que observamos questões referentes à redução de custos para os sistemas hospitalares e/ou de planos de saúde, não focando necessariamente no desfecho (prognóstico) de evolução para óbito e eventuais consequências advindas das ações tomadas estamos aplicando este tipo de pensamento. Um exemplo a ser dado seria o da identificação de pacientes cujo prognóstico é a evolução para óbito em 15, 30 ou 60 dias. Segundo [13], para situações de risco, a ética utilitarista é adequada "usando o utilitarismo de regras com exceção para avaliar a ação e suas consequências" e também coloca questões de bioética a serem consideradas para maximização de consenso sobre as normas. O utilitarismo não vai guiar a deliberação, mas dar as bases da justificativa da deliberação acerca de algum assunto.

Pensando, por outro lado, numa ética clássica, onde é examinada a intenção de quem executa a ação, pensando no que é virtude e o que é o bem, podemos dar como exemplo o caso de identificação de pacientes que tem prognóstico de evolução para óbito e então, a partir do julgamento médico a partir de evidências, decidir-se por utilizar medidas paliativas para cuidado do paciente ou não. Contudo, [14] adverte que a ética médica se separou da ética clássica em meados do século XX e que os direitos humanos passaram a tomar parte do que é a ética médica atual, dada a pluralidade de pensamentos e códigos de éticas nas mais diferentes partes do globo.

V. ENVOLVIMENTO DE PROFISSIONAIS

Por se tratar de um problema que envolve medicina, em especial cirurgias ortopédicas e internações, pessoas que potencialmente poderiam estar envolvidas com o problema seriam:

- cirurgiões com especialização em ortopedia
- pediatras ou geriatras, dependendo do público que vai realizar a cirurgia
- angiologistas uma vez que a operação, dependo da condição do paciente pode levar a um aspecto de trombose
- cardiologistas, dado que a pessoa que se submeterá à cirurgia pode ter problemas cardíacos

Tabela I

		Real		Total
		Positivo	Negativo	
Predito	Positivo	$TP(a)$	$FP(b)$	$a + b$
	Negativo	$FN(c)$	$TN(d)$	$c + d$
Total		$a + c$	$b + d$	N

Tabela II
MATRIZ DE CONFUSÃO.

- anestesistas que participam de cirurgias com especialização em ortopedia
- médicos intensivistas que atendem unidades de terapia intensiva
- enfermeiros que atendem alas cirúrgicas e de terapia intensiva
- médicos paliativistas que apoiam com cuidados paliativos os pacientes atendidos em unidades de terapia intensiva
- fisioterapeutas
- profissionais de terapia ocupacional
- profissionais que cuidam de suprimentos médicos nas áreas cirúrgicas e de terapia intensiva
- profissionais que trabalham em áreas de gestão financeira dos hospitais e planos de saúde envolvidos com custos de equipamento, estadia e remuneração de profissionais de saúde e suprimentos hospitalares.

VI. AVALIAÇÃO

A partir da Tabela I os aspectos que mais nos interessam são os que fazem com que os erros sejam minimizados (poucos falso positivos e poucos falsos negativos). Ainda assim, será preciso avaliar quais deles farão mais sentido no decorrer da evolução do trabalho, mas para um primeiro momento, podemos indicar que *precision*, *recall* e *acurácia (accuracy)* deveriam estar o mais próximo de 1 possível, uma vez que:

$$Precision = \frac{TP}{TP + FP} = \frac{a}{a + b}$$

$$Recall = \frac{TP}{TP + FN} = \frac{a}{a + c}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Conforme [15], vários aspectos clínicos que envolvem risco em cirurgias ortopédicas podem se colocar, como cardíaco, circulatório, trombose e tromboembolismo, alterações renais ou mesmo transfusões de sangue. Sobre internações pediátricas, [16] traumatismo crânio encefálico, lesões ortopédicas graves e politraumatismos fazem com que unidades de terapia pediátricas fiquem lotadas.

Como descreve [17], os sistemas de IA implantados emergem como um auxílio ao profissional, pois o médico possui autonomia na relação médico-paciente. Sendo assim, um dos cuidados é entender que o médico e paciente são os agentes principais e a inteligência artificial é acessória aos prognósticos, apoiando as decisões.

Outro ponto importante é deixar claro o que será utilizado para compor o resultado da predição. Neste sentido, uma inteligência artificial que tenha um nível adequado de explicabilidade pode ser interessante. [18] diz que a explicabilidade é essencial para que os usuários entendam, confiem e gerenciem aplicativos poderosos de inteligência artificial.

É importante também, do ponto de vista de trazer maior clareza a médicos e pacientes, explicitar a privacidade do sistema, garantindo e informando aos interessados que a estruturação do sistema, o uso de modelos e as respostas dadas por ele estão em conformidade com a LGPD [4] e que as normas de segurança de dados vigente são cumpridas.

Por outro lado, a fim de avaliar a segurança do sistema podemos empiricamente criar exemplos de dados adversários, alimentando-os e calculando a frequência em que o objetivo adversário consegue ser alcançado (redução de confiança, classificação incorreta, classificação incorreta direcionada ou classificação incorreta de origem/destino). Isso é possível pois sabemos quais dados de entrada contêm uma perturbação adversária e quais não.

Em contraste, uma forma de caracterizar a robustez adversarial é a pontuação CLEVER (cross-Lipschitz extreme value for network robustez), a pontuação (indiretamente) analisa a distância de um ponto de dados do sistema até o limite de decisão do classificador. Quanto maior a pontuação, mais robusto é o modelo. Podemos definir um valor mínimo aceitável para a pontuação média do CLEVER. Se o modelo atingir isso, podemos certificar com confiança um nível de segurança e robustez.

VII. PLANEJAMENTO

Para a execução dos passos de mineração e aprendizado de máquina, utilizaremos a metodologia presente na Figura 1 de [3] baseado no CRISP-DM [19].

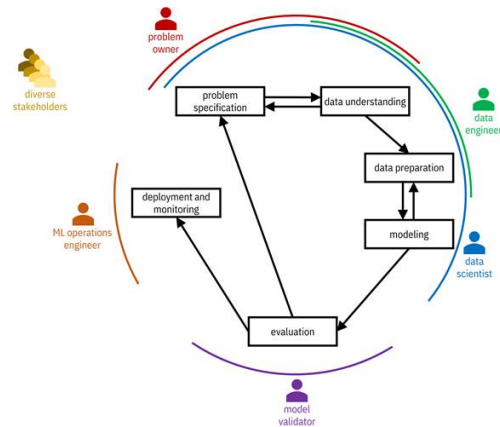


Figura 1. Etapas do ciclo de vida de um projeto de aprendizado de máquina conforme [3]

De forma sintética, as etapas do processo são as seguintes:

- Especificação do problema: o foco nesta etapa é entender a demanda a partir de uma perspectiva de negócios, afim de traçar estratégias para atingir os objetivos do problema.

- Entendimento dos Dados: avaliação, reconhecimento e caracterização do problema de dados com os dados propriamente ditos, além da identificação de conjuntos de interesses.
- Preparação dos Dados: construção do conjunto de dados a ser utilizado (sanitizações em data lake ou construções de data marts) a partir dos dados identificados na etapa anterior.
- Modelagem: construção de modelo estatístico, calibração de parâmetros (*fine tuning*) e otimizações.
- Avaliação: avaliação se o modelo estatístico ajuda a atingir e responder aos objetivos do problema de negócio traçado na etapa de especificação do problema.
- Implantação: entrega do resultado final para que este possa ser útil ao solicitante.

Para a etapa de modelagem do processo, [3] nos apresenta a Figura 2 com etapas de pré-processamento a partir do conjunto de dados de treino, em seguida o treinamento do modelo, a partir dos dados pré-processados, seguido de um pós processamento a partir da modelagem inicial e ao fim um modelo final é gerado.

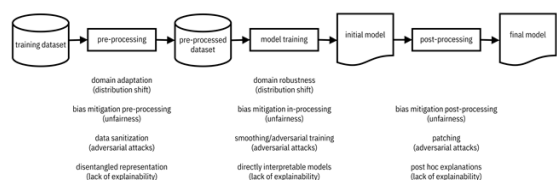


Figura 2. Principais componentes da etapa de modelagem estatística de um projeto de aprendizado de máquina conforme [3]

Mudança de distribuição, injustiça, ataques adversariais e ausência de explicação podem ser mitigados usando as várias técnicas listadas abaixo de cada etapa.

VIII. LIMITAÇÕES

As limitações para o presente trabalho são expostas abaixo:

- eventuais *outliers* poderão não ser tratados da forma adequada
- eventuais dados faltantes (*missing data*) poderão não ser tratados da forma adequada
- os dados foram entregues de forma anonimizada, de forma que não é possível ter clareza se a amostra é representativa da população
- o cruzamento dos dados das tabelas existente será feito com base em critérios técnicos, podendo não ser o mais adequado a ser executado
- as características dos dados serão extraídas considerando critérios técnicos. Isso pode fazer com que critérios médicos mais relevantes sejam colocados de lado.
- as pessoas que criaram o presente documento não tem domínio completo do tema para realizar qualquer julgamento sobre qualidade dos dados.

IX. POSITIONALITY STATEMENT

A aluna Bruna tem 25 anos, cresceu em Dracena no interior de São Paulo, frequentou apenas escolas públicas, sua graduação foi em Engenharia de Computação na UTFPR de Cornélio Procopio - PR. Atualmente é Cientista de Dados pleno no SiDi - Campinas, SP. Ela acredita que a inteligência artificial não substitui o conhecimento humano, mas possibilita um suporte maior para a tomada de decisão, tornando-a mais rápida e mais robusta, visto que consegue detectar padrões invisíveis muitas vezes ao olho nu. Entretanto, ainda há muitos aspectos para serem melhorados para que não reflitam os pré-conceitos existentes da sociedade no modelo, pois não acredita que seja justo que as pessoas estejam limitadas às limitações impostas pela sociedade para o grupo social pertencente.

O aluno Fabio frequentou escola particular no ensino fundamental e no ensino médio realizou-o parte em escola pública estadual, parte em escola particular em Campinas - SP. Sua graduação foi em Matemática na Universidade Estadual de Campinas — UNICAMP. Atualmente trabalha na CI&T, com sede também em Campinas, no papel de Cientista de Dados. Ele acredita que a inteligência artificial é uma tecnologia acessória ao conhecimento humano, ou seja, ela não substitui o conhecimento, mas acrescenta uma camada a mais a ele. Ainda assim, acredita que hoje a inteligência artificial acentua desigualdades, pois não vê, no âmbito social, o emprego massivo desta tecnologia para o bem comum. Entende que toda a pessoa é dotada de dignidade e que a Ciência deve apoiar o conhecimento humano em prol do bem comum, privilegiando os menos favorecidos afim de garantir segurança plena para todos (alimentar, de saúde, educacional, social, psicológica, entre outras).

Possivelmente o trabalho pode ter vieses do ponto de vista de entendimento de homem e de mundo focado num contexto de locais mais abastados (uma vez que ambos os alunos vieram do estado de São Paulo, um estado com IDH alto [20], comparativamente a outros estados brasileiros).

Outro aspecto importante e que pode enviesar o resultado é a característica de ambos de pensamento analítico, na medida em que ambos trabalham com ciências exatas. O conhecimento de áreas médicas é relativamente novo, de modo que potenciais decisões tomadas terão um aspecto que pode não ser verdade do ponto de vista da prática médica uma vez que nenhum dos dois é detentor de um conhecimento técnico-científico da área para emitir um posicionamento consistentemente embasado numa literatura médica robusta.

REFERÊNCIAS

- [1] “Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms.”
- [2] “Why it’s totally unsurprising that amazon’s recruitment ai was biased against women.”
- [3] K. R. Varshney, *Trustworthy Machine Learning*. Chappaqua, NY, USA: Independently Published, 2022.
- [4] TIVIT, “O papel das leis de proteção de dados no desenvolvimento de inteligência artificial.”
- [5] J. C. Patino, Cecilia Maria e Ferreira, “Prognostic studies for health care decision making,” 2019.
- [6] RCCC-eu, “Simplified acute physiology score iii,” 2022.

- [7] MDApp, “Palliative prognostic index (ppi).” <https://www.mdapp.co/palliative-prognostic-index-ppi-calculator-402/>, 2022. [Online; acessado 21-set-2022].
- [8] MDApp, “Palliative prognostic score (pap).” <https://www.mdapp.co/palliative-prognostic-score-pap-calculator-401/>, 2022. [Online; acessado 21-set-2022].
- [9] MD+Calc, “Maggic risk calculator for heart failure.” <https://www.mdcalc.com/maggic-risk-calculator-heart-failure>, 2022. [Online; acessado 21-set-2022].
- [10] E. W. Steyerberg, K. G. M. Moons, D. A. van der Windt, J. A. Hayden, P. Perel, S. Schroter, R. D. Riley, H. Hemingway, D. G. Altman, and for the PROGRESS Group, “Prognosis research strategy (progress) 3: Prognostic model research,” *PLoS Medicine*, vol. 10, pp. 1–9, 02 2013.
- [11] E. Toreini, M. Aitken, K. Coopamootoo, K. Elliott, C. G. Zelaya, and A. van Moorsel, “The relationship between trust in ai and trustworthy machine learning technologies,” 2019.
- [12] R. Lopes, “Há espaço para uma concepção não moral da normatividade prática em nietzsche?: notas sobre um debate em andamento,” *Cadernos Nietzsche*, pp. 89–134, 2013.
- [13] D. A. das Neves, “O critério utilitarista será adequado para situação de risco?,” *Revista Brasileira de Saúde Materno Infantil*, vol. 10, pp. s347–s353, Dec. 2010.
- [14] T. H. B. Sanchez and I. C. Fraiz, “Ética médica e formação do médico,” *Revista Bioética*, vol. 30, pp. 284–299, June 2022.
- [15] L. E. G. Leme, M. do Carmo Sitta, M. Toledo, and S. da Silva Henriques, “Cirurgia ortopédica em idosos: aspectos clínicos,” *Revista Brasileira de Ortopedia*, vol. 46, no. 3, pp. 238–246, 2011.
- [16] N. O. W. Batista, M. C. d. R. Coelho, S. M. Trugilho, G. C. Pinasco, E. F. d. S. Santos, and V. Ramos-Silva, “Clinical-epidemiological profile of hospitalised patients in paediatric intensive care unit,” *Journal of Human Growth and Development*, vol. 25, pp. 187 – 193, 00 2015.
- [17] B. Q. Gama, M. L. F. S. Alves, I. F. Dos Reis, V. L. Queiroga, M. E. F. Silva, J. de Paula Silva, and J. R. Alves, “A relevância da inteligência artificial no cenário da cardiologia,” *Revista Atenas Higeia*, vol. 4, no. 2, 2022.
- [18] D. Gunning, M. Stefik, J. Choi, T. Miller, S. Stumpf, and G.-Z. Yang, “Xai—explainable artificial intelligence,” *Science robotics*, vol. 4, no. 37, p. eaay7120, 2019.
- [19] C. Shearer, “The crisp-dm model: the new blueprint for data mining,” *Journal of data warehousing*, vol. 5, no. 4, pp. 13–22, 2000.
- [20] IBGE, “Índice de desenvolvimento humano.” <https://cidades.ibge.gov.br/brasil/sp/pesquisa/37/0?tipo=ranking&ano=2010>, 2022. [Online; acessado 23-set-2022].