

# 1 Introdução

Diversos modelos de psicoterapia apresentam evidência de eficácia no tratamento de transtornos mentais. Uma parcela significativa dos pacientes, no entanto, não responde às intervenções, podendo até mesmo apresentar piora quando submetida a tratamento psicoterápico [Cuijpers2021]. A variabilidade na taxa de resposta ao tratamento é parcialmente explicada pelo quadro clínico do paciente; ainda assim, observou-se que pessoas de uma mesma população clínica respondem de maneiras diferentes a um mesmo tratamento [Hofmann2012]. Especula-se que um atendimento personalizado a nível individual possa melhorar a taxa de resposta ao tratamento [Norcross2010, Norcross2018].

A literatura aponta uma série de variáveis preditoras de desfecho clínico para pacientes em tratamento psicoterápico, o que permitiria a adaptação dos protocolos de intervenção de acordo com a expectativa de resposta [Smagula2019, Andover2020]. Não existe, porém, consenso acerca do valor preditivo de cada variável a nível individual e interações complexas entre diferentes preditores podem prejudicar a acurácia das previsões e impactar propostas de personalização de tratamento negativamente [Taubitz2022].

Técnicas de inteligência artificial baseadas em aprendizagem de máquina apresentam a capacidade de integrar uma grande quantidade de dados, impondo poucas restrições ao comportamento das variáveis observadas e produzindo modelos flexíveis aplicáveis em diferentes contextos [Dwyer2018]. Em essência, a aprendizagem de máquina consiste no uso de métodos estatísticos e computacionais para identificar padrões de relacionamento subjacentes a um grande conjunto de dados, permitindo a construção de modelos discriminativos ou generativos [Roth2018]. Estudos sobre a aplicação de modelos de aprendizagem de máquina na previsão do desfecho de tratamentos em saúde mental apresentam resultados promissores [Dwyer2018], o que possibilitaria maior assertividade na personalização de tratamentos psicoterápicos a nível individual.

Embora representem um potencial ganho para a prática em psicologia clínica, as tecnologias de aprendizagem de máquina em saúde mental não são apresentadas durante a formação de profissionais da psicologia. As diretrizes nacionais curriculares para cursos de graduação em psicologia propõe um modelo formativo generalista, deixando pouco espaço para a apresentação de conhecimentos metodológicos muito específicos como a aprendizagem de máquina [CNE2023, Ruda2019].

Este trabalho tem por objetivo apresentar os conceitos básicos relacionados à aprendizagem de máquina, de modo que possa servir como uma introdução acessível ao tema. Busca-se também ilustrar o uso da aprendizagem de máquina no auxílio à tomada de decisões para o planejamento de intervenções psicoterápicas. Utilizando um conjunto de dados de uma intervenção digital em psicologia positiva para depressão [Collins2023], pretende-se construir um modelo de árvore de decisão para prever o desfecho da intervenção para novos pacientes.

## 2 Conceitos Básicos de Aprendizagem de Máquina

### 2.1 O que é aprendizagem de máquina?

Aprendizagem de máquina é a área da ciência da computação que tem como objetivo geral o desenvolvimento de programas de computador capazes de aprender a realizar uma tarefa sem serem explicitamente programados [Bi2019, Theobald2021]. Neste contexto, aprendizagem refere-se a aplicação de procedimentos estatísticos e computacionais sobre um conjunto de informações empíricas, buscando alcançar melhorias de desempenho em uma determinada tarefa [Theobald2021].

Aprender trata-se, portanto, de ajustar os parâmetros de um modelo estatístico e computacional aos dados observados de modo a maximizar o desempenho na tarefa em questão [Bi2019]. Programas de computador baseados em aprendizagem de máquina são capazes de identificar padrões de interação complexos entre variáveis em conjuntos de dados com alta dimensionalidade para realizar tarefas de classificação, regressão, agrupamento e outras [Theobald2021].

Considere, por exemplo, um estudo observacional hipotético que investiga a relação entre características de personalidade e o nível de satisfação profissional entre psicólogos. O estudo baseia-se no modelo dos cinco grandes fatores da personalidade [Hutz2018] e usa o instrumento da Bateria Fatorial da Personalidade para coleta de dados, registrando as pontuações obtidas nas escalas de neuroticismo, extroversão, socialização, realização e abertura [Sancineto2015]. Além disso, os participantes do estudo reportam o próprio nível de satisfação profissional em uma escala que contém os seguintes valores: baixo, médio e alto. O conjunto de dados coletados é como apresentado na tabela 1.

neuroticismo	extroversão	socialização	realização	abertura	satisfação
1	1	2	4	5	alto
1	1	2	4	5	baixo
1	1	2	4	5	médio
1	1	2	4	5	médio

Table 1: Exemplo de dados coletados no estudo hipotético.

É possível utilizar esse conjunto de dados para construir um modelo de aprendizagem de máquina preditivo. Um algoritmo processa o conjunto de dados, identificando os padrões de interação existentes entre as variáveis preditoras (características de personalidade) e o desfecho de interesse (nível de satisfação profissional). O conhecimento adquirido durante o processamento dos dados é codificado nos parâmetros de um modelo de aprendizagem de máquina. O modelo pode então ser utilizado para fazer previsões sobre o nível de satisfação profissional de um indivíduo qualquer a partir de suas características de personalidade.

## 2.2 Os tipos de aprendizagem de máquina

As técnicas de aprendizagem de máquina podem ser organizadas de diferentes maneiras, incluindo classificação pela estratégia adotada durante o processo de aprendizagem e pelo objetivo geral de aprendizagem [Theobald2021, Ng2001].

### 2.2.1 Aprendizagem supervisionada, não supervisionada e por reforço

As categorias mais comumente usadas na descrição de modelos de aprendizagem de máquina dizem respeito à estratégia de aprendizagem adotada. O modelo pode ser construído segundo uma abordagem de aprendizagem supervisionada, aprendizagem não supervisionada ou aprendizagem por reforço [Theobald2021, Bi2019].

A aprendizagem supervisionada assemelha-se ao processo de aprendizagem adotado por seres humanos, onde o aprendiz identifica padrões a partir de um conjunto de exemplos preparado por um tutor. Durante a fase de aprendizagem, o modelo é exposto a um conjunto de dados que contém informações sobre o desfecho de interesse para cada uma das observações. O acesso às informações de desfecho providas por um agente externo confere o caráter de supervisão a este processo. Técnicas de aprendizagem de máquina para regressão e classificação (support vector machines, árvores de decisão, redes neurais) pertencem a esta categoria [Theobald2021, Bi2019]. Um exemplo para a aplicação deste tipo de aprendizagem é usar de dados de ensaios clínicos, onde o desfecho para cada paciente é conhecido, na construção de um modelo capaz de prever o resultado da intervenção para novos pacientes [Collins2023].

Na aprendizagem não supervisionada, o conjunto de dados analisado não contém qualquer informações sobre desfecho de interesse. Espera-se que o modelo identifique os padrões de relacionamento existentes entre as variáveis do conjunto de dados e gere agrupamentos ou projeções de maneira autônoma. Técnicas de aprendizagem de máquina para tarefas de agrupamento e redução de dimensionalidade (k-means clustering, PCA, TSNE) pertencem a esta categoria [Theobald2021, Bi2019]. Um exemplo para a aplicação deste tipo de aprendizagem é investigar os padrões de comorbidade em uma determinada população clínica [Sanchez2019].

Na aprendizagem por reforço, o modelo aprende através de repetidos ciclos de tentativa e erro. A cada ciclo de aprendizagem, o modelo recebe feedback sobre seu desempenho na tarefa, o feedback é incorporado à base de conhecimento construída pelo modelo em ciclos passados e, assim, melhora seu desempenho progressivamente [Theobald2021, Bi2019]. Um exemplo para a aplicação deste tipo de aprendizagem é auxiliar a tomada de decisões de tratamento em condições crônicas como a esquizofrenia [Shortreed2010].

### **2.2.2 Modelos discriminativos e generativos**

## **2.3 A construção de uma aplicação de machine learning**

### **2.3.1 Análise descritiva**

### **2.3.2 Pré-processamento**

Na etapa de pré-processamento, busca-se preparar o conjunto de dados de treinamento, colocando-no em um estado adequado à técnica de aprendizagem de máquina que se pretende utilizar. Tratamentos comumente realizados na etapa de pré-processamento são: seleção de características, transformações, imputações e balanceamento de classes.

A seleção de características consiste em eliminar do conjunto de dados as variáveis que tenham pouca contribuição para a aprendizagem da tarefa. Em um conjunto de dados onde todas as observações são de pessoas brasileiras, a variável de nacionalidade não contribui para a explicação do desfecho que se busca prever, portanto pode ser removida.

Transformações são aplicadas de acordo com os requisitos da técnica de aprendizagem de máquina em uso. Por exemplo, algumas técnicas de aprendizagem de máquina são suscetíveis à influência de variáveis com escala muito superior às demais; nesses casos é comum transformação das variáveis para uma escala padronizada (medida em desvios padrão a partir da média).

### **2.3.3 Treinamento do modelo**

### **2.3.4 Validação do modelo**

## **2.4 Parte 2**

Texto da parte 2.

### **3 Conclusão**

Texto da conclusão.