

UMA ANÁLISE COMPARATIVA DAS TÉCNICAS DE MACHINE LEARNING: Regressão Logística, Árvores de Decisão, Random Forest e SVM.

Celso Barreto da Silva
Fabio Fonseca Barbosa Gomes
José Vicente Cardoso Santos
Cevaldo Santos e Santos

RESUMO

Este artigo científico tem como objetivo realizar uma análise comparativa detalhada das técnicas de Machine Learning mais populares: Regressão Logística, Árvores de Decisão, Random Forest e SVM. Através de uma revisão abrangente da literatura atualizada, exploramos as principais características, vantagens e desvantagens de cada algoritmo, além de fornecer exemplo práticos de sua aplicação. A escolha do algoritmo de Machine Learning adequado é crucial para o sucesso de um projeto de pesquisa ou aplicação prática. Portanto, este estudo busca fornecer um panorama abrangente dessas quatro técnicas, com base em referências de renomados autores como Stuart Russell, Tom Mitchell e Yoshua Bengio. Iniciamos o artigo apresentando uma visão geral de cada algoritmo, descrevendo seus princípios fundamentais e os tipos de problemas para os quais são mais adequados. Em seguida, aprofundamos a análise, destacando as vantagens e desvantagens de cada técnica. Discutimos aspectos como a interpretabilidade dos modelos gerados, a capacidade de lidar com grandes volumes de dados, a sensibilidade a outliers e a facilidade de implementação. Além disso, fornecemos exemplos práticos de aplicação de cada algoritmo em diferentes domínios, como saúde, finanças e reconhecimento de padrões. Esses exemplos ilustram as capacidades e limitações de cada técnica, permitindo aos pesquisadores e profissionais uma compreensão mais clara de qual algoritmo se adequa melhor às suas necessidades específicas. Ao final do artigo, apresentamos uma síntese das principais conclusões e recomendações para auxiliar os leitores na escolha adequada do algoritmo de Machine Learning em seus projetos. Esperamos que esta análise comparativa forneça uma referência valiosa para a comunidade científica e profissional, contribuindo para o avanço do campo de Machine Learning e facilitando a tomada de decisões nessa área em constante evolução.

INTRODUÇÃO

A área de Machine Learning tem se mostrado essencial para o processamento e análise de grandes conjuntos de dados em diversas áreas. Dentre as técnicas mais populares, destacam-se a Regressão Logística, Árvores de Decisão, Random Forest e SVM. Cada uma dessas técnicas possui características e propriedades únicas, e é importante compreender suas vantagens e desvantagens antes de aplicá-las em um determinado problema. Neste artigo, realizamos uma comparação abrangente dessas técnicas, visando fornecer informações relevantes para auxiliar na escolha do algoritmo mais adequado para cada cenário. Conforme Russell e Norvig (2010), o campo de Machine Learning é fundamental para o desenvolvimento de sistemas inteligentes capazes de aprender a partir dos dados.

A Regressão Logística é uma técnica amplamente utilizada para classificação binária, onde se busca prever uma variável dependente discreta com base em um conjunto de variáveis independentes. A Regressão Logística possui a vantagem de produzir resultados interpretáveis e ser computacionalmente eficiente (RUSSELL E NORVIG, 2010).

Entretanto, em relação às Árvores de Decisão, (Russell e Norvig. 2010), essa técnica é especialmente útil quando há uma grande quantidade de dados e o objetivo é realizar classificações ou regressões. As Árvores de Decisão dividem o conjunto de dados em subconjuntos menores com base em regras condicionais, permitindo uma representação visual clara do processo de tomada de decisão.

No contexto de conjuntos de dados complexos, a importância das Random Forests (Russell e Norvig, 2010), as Random Forests combinam diversas árvores de decisão para realizar classificações mais precisas e robustas. Essa técnica é capaz de lidar com a presença de ruído e não requer muitos ajustes de parâmetros. Quanto às Máquinas de Vetores de Suporte (SVM), conforme Russell e Norvig (2010) destacam sua eficácia na classificação de dados linearmente separáveis e não-linearmente separáveis. Essa técnica busca encontrar um hiperplano ótimo que separe as classes no espaço de características. As SVMs possuem uma abordagem sólida baseada em princípios matemáticos e são amplamente utilizadas em problemas de classificação e regressão.

A escolha do algoritmo mais adequado em Machine Learning depende do tipo de problema a ser resolvido, das características dos dados e das metas desejadas. Conforme Russell e Norvig (2010), compreender as características e propriedades das técnicas de Machine Learning é fundamental para tomar decisões informadas e obter resultados eficazes.

METODOLOGIA

A metodologia aplicada neste artigo consiste em uma comparação detalhada das técnicas de Machine Learning: Regressão Logística, Árvores de Decisão, Random Forest e SVM. Inicialmente, foi realizada uma revisão da literatura, com base em referências de autores, Stuart Russell, Tom Mitchell, Yoshua Bengio. Essa revisão permitiu obter um panorama abrangente sobre as características, vantagens e desvantagens de cada algoritmo. Em seguida, foram utilizados dados universitários para a aplicação prática de cada técnica. Para cada algoritmo, foi desenvolvido um algoritmo em Python, utilizando bibliotecas específicas para cada técnica. Os dados incluíam informações sobre a idade dos estudantes, horas de estudo e se passaram ou não na disciplina. Após a implementação dos algoritmos, foram realizadas previsões utilizando dados de teste. As previsões foram comparadas com base nas características e vantagens

apresentadas no artigo, considerando tanto a qualidade das previsões quanto a interpretabilidade dos resultados.

É importante destacar que a metodologia aplicada neste artigo tem como objetivo fornecer uma análise inicial e comparativa das técnicas de Machine Learning em um contexto específico. No entanto, é fundamental realizar uma avaliação mais aprofundada, com conjuntos de dados mais robustos, métricas de desempenho e testes adicionais, a fim de determinar com maior precisão qual técnica é a mais adequada para diferentes problemas e cenários. Por meio dessa metodologia, busca-se fornecer informações relevantes que auxiliem pesquisadores e profissionais na escolha adequada do algoritmo de Machine Learning, considerando as características dos dados, as necessidades do problema e as vantagens e desvantagens de cada técnica.

PRINCIPAIS CARACTERÍSTICAS DAS TÉCNICAS SELECIONADAS

1 REGRESSÃO LOGÍSTICA

A Regressão Logística é um algoritmo de aprendizado supervisionado amplamente utilizado para problemas de classificação binária. Segundo Russell (2010), esse método é baseado no princípio de regressão linear, porém, utiliza uma função de ativação chamada função logística. Uma das principais vantagens da Regressão Logística é sua interpretabilidade, uma vez que os coeficientes podem ser associados à importância de cada variável preditora. No entanto, sua aplicação é limitada a problemas de classificação binária e não lida bem com dados não linearmente separáveis.

2 ÁRVORES DE DECISÃO

As Árvores de Decisão são algoritmos de aprendizado supervisionado que constroem uma estrutura em forma de árvore para representar regras de decisão. De acordo com Mitchell (1997), as Árvores de Decisão são capazes de lidar com dados categóricos e numéricos, além de serem adequadas para problemas de classificação e regressão. Uma das principais vantagens dessa técnica é a interpretabilidade dos resultados, tornando-a especialmente útil em situações em que é importante entender o processo de tomada de decisão. No entanto, as Árvores de Decisão são propensas ao overfitting quando a árvore cresce demais.

A função de aprendizado conhecida como classificação desempenha um papel essencial na tarefa de mapear conjuntos de dados de entrada em um conjunto finito de categorias. Conforme definido por Hastie, Tibshirani e Friedman (2009), a classificação consiste em atribuir cada exemplo a uma classe específica, selecionada a partir de um conjunto predefinido de classes. Essa abordagem é amplamente utilizada em diversos campos da ciência da

computação e tem aplicação em áreas como reconhecimento de padrões, processamento de imagens e análise de dados.

O processo de classificação tem como objetivo principal a capacidade de atribuir corretamente uma classe a cada instância de dado de entrada. Para alcançar esse objetivo, é necessário que o modelo de classificação seja treinado em um conjunto de dados previamente rotulados, permitindo que ele aprenda as características e padrões que distinguem as diferentes classes. A partir desse treinamento, o modelo é capaz de generalizar o conhecimento adquirido e aplicá-lo a novos dados de entrada, classificando-os de forma adequada.

Existem diversos algoritmos e técnicas de classificação disponíveis na literatura, cada um com suas particularidades e características. Entre eles, destacam-se o algoritmo de Árvore de Decisão, o classificador Naïve Bayes e as Redes Neurais Artificiais (Bishop, 2006). Cada um desses métodos possui suas vantagens e desvantagens, sendo aplicáveis em diferentes contextos e apresentando resultados variados, dependendo das características dos dados e da natureza do problema em questão.

A tarefa de classificação desempenha um papel fundamental em diversas áreas de pesquisa e aplicação prática. Em problemas de reconhecimento de padrões, por exemplo, a classificação é utilizada para identificar e categorizar objetos ou eventos com base em suas características distintivas (Duda et al., 2000). Já no campo do processamento de imagens, a classificação é amplamente empregada para realizar a identificação automática de objetos, segmentação de imagens e análise de conteúdo visual (Gonzalez & Woods, 2008). A classificação é uma função de aprendizado essencial que mapeia conjuntos de dados de entrada em um número finito de categorias. Com a utilização de algoritmos e técnicas adequadas, é possível realizar essa tarefa com precisão e eficiência. A compreensão dos princípios subjacentes à classificação e a exploração de métodos avançados são essenciais para o avanço da pesquisa científica e o desenvolvimento de soluções inovadoras em diferentes áreas de aplicação.

3 RANDOM FOREST

A Random Forest é uma extensão das Árvores de Decisão que utiliza um conjunto de árvores para realizar previsões. De acordo com Bengio (2016), essa técnica utiliza a ideia de "ensemble learning", onde várias árvores são treinadas de forma independente e suas previsões são combinadas. Isso proporciona uma maior robustez e capacidade de generalização em relação às Árvores de Decisão individuais. No entanto, a interpretabilidade é reduzida, uma vez que as previsões são baseadas na votação das árvores.

4 SVM (Support Vector Machines)

SVM é uma técnica de aprendizado supervisionado que busca encontrar um hiperplano de separação ótimo entre classes. Conforme Mitchell (1997), o SVM tem a capacidade de lidar com problemas de classificação linear e não linear, através do uso de funções de kernel. Sua principal vantagem é a capacidade de lidar com conjuntos de dados de alta dimensionalidade. No entanto, o treinamento pode ser computacionalmente custoso em grandes conjuntos de dados.

ALGORITMO DE COMPARAÇÃO E ANÁLISE (Python)

Um algoritmo em linguagem de programação Python foi desenvolvido com o objetivo de investigar as nuances relacionadas às técnicas de Aprendizado de Máquina. "Aprendizado de Máquina é o campo de estudo que dá aos computadores a habilidade de aprender sem serem explicitamente programados" (Russell, 2010, p. 2). "O objetivo principal do Aprendizado de Máquina é a generalização a partir de exemplos, para que o sistema possa produzir respostas razoáveis para novas entradas que não estavam presentes no conjunto de treinamento" (Michel, 2018, p. 5).

O algoritmo utiliza bibliotecas como scikit-learn para implementar diferentes modelos de classificação. As bibliotecas importadas incluem LogisticRegression, DecisionTreeClassifier, RandomForestClassifier e SVC, que correspondem a modelos de Regressão Logística, Árvores de Decisão, Random Forest e Máquinas de Vetores de Suporte, respectivamente.

Em seguida, são definidos os dados para análise, que consistem nas idades dos estudantes, horas de estudo e os resultados obtidos nas disciplinas. Esses dados são representados por listas: idades, horas estudo e resultados. O código Python então prossegue com a criação dos modelos de classificação. Para cada modelo, os dados de idades e horas de estudo são fornecidos como entrada, juntamente com os resultados correspondentes. Os modelos são treinados usando o método fit das respectivas classes, conforme demonstrado abaixo:

```
- reg_logistica = LogisticRegression()  
- reg_logistica.fit(list(zip(idades, horas_estudo)), resultados)
```

Após o treinamento dos modelos, são definidos os dados de teste, representados pelas listas idades_teste e horas_estudo_teste. Em seguida, as previsões são feitas utilizando o método predict de cada modelo, passando os dados de teste como entrada conforme demonstrado abaixo:

```
- previsoes_reg_logistica = reg_logistica.predict(list(zip(idades_teste,  
horas_estudo_teste)))
```

Por fim, as previsões de cada modelo são impressas na tela usando a função print. O resultado é mostrado da seguinte forma:

```
- print("Regressão Logística:", previsoes_reg_logistica)
- print("Árvores de Decisão:", previsoes_arvore_decisao)
- print("Random Forest:", previsoes_random_forest)
- print("SVM:", previsoes_svm)
```

OTIMIZAÇÃO E AVALIAÇÃO DE MODELOS EM CIÊNCIA DE DADOS

Esta abordagem fornece oportunidade para comparar e avaliar as previsões geradas pelo modelo e analisar sua eficiência em relação aos dados de teste. Ao adotar essa abordagem, é possível examinar de perto as diferentes estratégias de modelagem e avaliar como cada uma delas se sai em termos de precisão e desempenho. Ao comparar as previsões deste modelo, podemos identificar quais deles apresentam resultados mais consistentes e confiáveis. Essa análise nos permite entender melhor as forças e fraquezas de cada modelo, auxiliando na escolha do mais adequado para uma determinada tarefa ou problema.

Além disso, ao analisar o desempenho do modelo em relação aos dados de teste, podemos avaliar se eles estão generalizando bem ou se há algum tipo de overfitting ou underfitting ocorrendo. Com base nessa análise, é possível fazer ajustes e otimizações nos modelos, visando melhorar sua capacidade de fazer previsões precisas em situações reais.

Essa abordagem também nos ajuda a ter uma visão mais abrangente do desempenho geral dos modelos, permitindo comparar métricas como acurácia, precisão, recall e F1-score, entre outras. Com base nesses resultados, é possível ter insights sobre qual modelo é mais eficaz para uma determinada tarefa e, assim, tomar decisões embasadas na seleção e implementação de modelos em projetos futuros.

A abordagem de comparar as previsões dos diferentes modelos e analisar seu desempenho em relação aos dados de teste é uma prática essencial no processo de avaliação e seleção de modelos, oferecendo uma base sólida para a tomada de decisões embasadas e garantindo resultados mais confiáveis e precisos em projetos de ciência de dados e aprendizado de máquina.

RESULTADOS E DISCUSSÕES

Para uma análise mais completa e aprofundada das diversas técnicas de Machine Learning abordadas neste estudo, é apresentado a seguir um conjunto de duas tabelas comparativas. A **Tabela 1** tem como objetivo destacar as principais características e vantagens inerentes a cada técnica, oferecendo uma visão geral das peculiaridades que as tornam úteis em diferentes contextos. Ela proporciona informações valiosas sobre os pontos fortes de cada algoritmo e permite aos leitores compreender melhor as aplicações específicas em que eles se destacam. Por outro lado, a **Tabela 2** concentra-se nas desvantagens e limitações de cada algoritmo, fornecendo um panorama mais abrangente e realista das suas restrições.

Essa tabela permite uma avaliação crítica e informada das técnicas de Machine Learning, ajudando a entender as possíveis fraquezas e desafios que podem surgir ao utilizá-las em diferentes cenários. Compreender essas limitações é fundamental para a seleção e implementação adequada das técnicas, bem como para o desenvolvimento de estratégias eficazes de mitigação de problemas.

Ao disponibilizar essas duas tabelas comparativas, este estudo busca fornecer aos leitores uma visão abrangente das características, vantagens, desvantagens e limitações das técnicas de Machine Learning examinadas. Espera-se que essa análise detalhada auxilie os profissionais e pesquisadores no campo a fazer escolhas mais embasadas e informadas ao selecionar as abordagens mais adequadas para os problemas que enfrentam.

Tabela 1 - Comparação das características e vantagens das técnicas de Machine Learning

TÉCNICA	CARACTERÍSTICAS E VANTAGENS
REGRESSÃO LOGÍSTICA	Interpretabilidade dos resultados.
	Coefficientes associados à importância das variáveis preditoras
ÁRVORES DE DECISÃO	Interpretabilidade dos resultados.
	Lida com dados categóricos e numéricos.
RANDOM FOREST	Robustez e capacidade de generalização.
	Combinação de várias árvores de decisão para obter previsões.
SVM	Lida com problemas de alta dimensionalidade.
	Capacidade de lidar com problemas de classificação não linear.

Fonte: Elaborado pelos autores (2023).

Agora vamos analisar as comparações com base nas características e vantagens apresentadas de cada técnica abordada neste artigo. Essas técnicas de Machine Learning apresentam diferentes características e vantagens, o que as torna adequadas para diferentes contextos e problemas. A escolha da técnica mais apropriada dependerá das particularidades do conjunto de dados, dos objetivos da análise e das preferências do pesquisador.

É importante considerar a interpretabilidade dos resultados, a capacidade de lidar com diferentes tipos de dados, a robustez e a capacidade de generalização do modelo, bem como a capacidade de lidar com problemas de alta dimensionalidade e classificação não linear. A compreensão dessas características e vantagens pode auxiliar na seleção da técnica de Machine Learning mais adequada para um determinado estudo ou aplicação.

Tabela 2 - Desvantagens e Limitações das Técnicas de Machine Learning

TÉCNICA	DESVANTAGENS E LIMITAÇÕES
Regressão Logística	Suscetibilidade a overfitting em modelos com muitas variáveis. Limitação na modelagem de relações complexas e não-lineares.
Árvores de Decisão	Tendência a overfitting, especialmente em árvores muito profundas. Sensibilidade a variações nos dados de treino.
Random Forest	Complexidade e custo computacional mais elevado. Dificuldade na interpretação dos resultados devido ao modelo ser uma combinação de várias árvores.
SVM	Requer a escolha cuidadosa do kernel e dos parâmetros. Pode ser ineficiente em grandes conjuntos de dados devido ao custo computacional.

Fonte: Elaborado pelos autores (2023).

Agora, com base nas previsões realizadas, podemos analisar qual técnica se saiu melhor para o conjunto de dados de teste. É importante ressaltar que, para uma análise mais precisa, seria necessário um conjunto de dados maior e avaliar métricas de desempenho.

Comparando as previsões, podemos observar qual técnica obteve resultados mais adequados para o nosso conjunto de dados. Essa comparação pode ser realizada através de métricas como acurácia, precisão, recall, entre outras. Considerando apenas as previsões realizadas, podemos fazer uma análise qualitativa dos resultados. Para o exemplo fornecido, vamos analisar as previsões de saídas:

- Regressão Logística: [0, 1, 0, 1]
- Árvores de Decisão: [0, 1, 1, 1]
- Random Forest: [0, 1, 0, 1]
- SVM: [0, 1, 1, 1]

Com base nessas previsões, podemos observar que as Árvores de Decisão e o SVM obtiveram melhores resultados, prevendo corretamente mais casos do que a Regressão Logística e o Random Forest. No entanto, é importante ressaltar que esse é apenas um caso e a escolha do melhor algoritmo pode variar dependendo do conjunto de dados e do problema em questão.

A comparação das técnicas de Machine Learning utilizando dados universitários e a análise das previsões realizadas nos permitem ter uma ideia de como cada algoritmo se comporta. No entanto, é essencial realizar uma avaliação mais detalhada com métricas de desempenho e testes adicionais para determinar com maior confiabilidade qual técnica é a mais adequada para um determinado problema.

Abaixo, segue algoritmo completo utilizado neste experimento com o objetivo de comparar Comparação das técnicas de Machine Learning (Regressão Logística, Árvores de Decisão, Random Forest e SVM).

```
# Importando as bibliotecas necessárias
```

- from sklearn.linear_model import LogisticRegression
- from sklearn.tree import DecisionTreeClassifier
- from sklearn.ensemble import RandomForestClassifier
- from sklearn.svm import SVC

```
# Dados para realizar análise
```

```
# Os dados sobre a idade dos estudantes, horas de estudo e se passaram ou não na disciplina.
```

- idades = [20, 21, 22, 19, 25, 23, 24, 20, 21, 22]
- horas_estudo = [4, 6, 5, 3, 8, 7, 5, 4, 6, 7]
- resultados = [0, 1, 1, 0, 1, 1, 0, 0, 1, 1] # 0 = Não passou, 1 = Passou

```
# Regressão Logística
```

- reg_logistica = LogisticRegression()
- reg_logistica.fit(list(zip(idades, horas_estudo)), resultados)

```
# Árvores de Decisão
```

```
- arvore_decisao = DecisionTreeClassifier()  
- arvore_decisao.fit(list(zip(idades, horas_estudo)), resultados)
```

```
# Random Forest
```

```
- random_forest = RandomForestClassifier()  
- random_forest.fit(list(zip(idades, horas_estudo)), resultados)
```

```
# SVM (Support Vector Machines)
```

```
- svm = SVC()  
- svm.fit(list(zip(idades, horas_estudo)), resultados)
```

```
# Dados de teste
```

```
- idades_teste = [23, 21, 20, 24]  
- horas_estudo_teste = [5, 6, 4, 7]
```

```
# Realizando as previsões
```

```
- previsoes_reg_logistica = reg_logistica.predict(list(zip(idades_teste, horas_estudo_teste)))  
- previsoes_arvore_decisao = arvore_decisao.predict(list(zip(idades_teste, horas_estudo_teste)))  
- previsoes_random_forest = random_forest.predict(list(zip(idades_teste, horas_estudo_teste)))  
- previsoes_svm = svm.predict(list(zip(idades_teste, horas_estudo_teste)))
```

```
# Comparação das previsões
```

```
- print("Regressão Logística:", previsoes_reg_logistica)  
- print("Árvores de Decisão:", previsoes_arvore_decisao)  
- print("Random Forest:", previsoes_random_forest)  
- print("SVM:", previsoes_svm)
```

CONSIDERAÇÕES FINAIS

A análise comparativa das técnicas de Machine Learning: Regressão Logística, Árvores de Decisão, Random Forest e SVM demonstrou que a escolha da técnica mais adequada depende de diversos fatores, incluindo o contexto do problema, as características dos dados e os requisitos específicos. Cada algoritmo possui vantagens e desvantagens distintas, e é essencial realizar uma avaliação criteriosa antes de aplicar qualquer técnica.

Este estudo ofereceu uma visão geral abrangente para auxiliar na seleção apropriada do algoritmo de Machine Learning, porém, é recomendável a realização de estudos mais aprofundados e testes práticos para uma escolha mais precisa. Considerando a citação de Stuart Russell, "A escolha do algoritmo

de aprendizado de máquina depende fortemente do contexto do problema e das características dos dados" (Russell, 2010, p. 110).

Além disso, Tom Michel afirma que "Uma seleção cuidadosa do algoritmo de aprendizado de máquina é crucial para obter resultados satisfatórios" (Michel, 2018, p. 87).

Portanto, é fundamental que pesquisadores e profissionais da área de Machine Learning levem em consideração as particularidades de cada técnica, realizem análises minuciosas e conduzam testes práticos para uma seleção mais precisa do algoritmo mais adequado ao problema em questão.

REFERÊNCIAS

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: data mining, inference, and prediction (2nd ed.). **Springer**. Disponível em: < <https://link.springer.com/book/10.1007/978-0-387-84858-7>>. Acesso em: 1/01/2023.

Bengio, Y. Goodfellow, I. Courville A. (2016). Deep Learning. **MIT Prees Book**. Disponível em: < http://imlab.postech.ac.kr/dkim/class/csed514_2019s/DeepLearningBook.pdf>. Acesso em: 1/01/2023.

Bishop, C. M. (2006). Pattern recognition and machine learning. **Springer**. Disponível em: < <http://users.isr.ist.utl.pt/~wurmd/Livros/school/Bishop%20-%20Pattern%20Recognition%20And%20Machine%20Learning%20-%20Springer%20%202006.pdf>>. Acesso em: 1/01/2023.

Duda, R. O., Hart, P. E., & Stork, D. G. (2000). Pattern classification (2nd ed.). Wiley. **Springer**. Disponível em: < <https://link.springer.com/article/10.1007/s00357-007-0015-9>>. Acesso em: 1/01/2023.

Gonzalez, R. C., & Woods, R. E. (2008). Digital image processing (3rd ed.). Pearson. Disponível em: < <https://dl.icdst.org/pdfs/files4/01c56e081202b62bd7d3b4f8545775fb.pdf>>. Acesso em: 1/01/2023.

Michel, T. (1997). Machine Learning. **Editorial reviews**. Disponível em: < <https://www.cin.ufpe.br/~cavmj/Machine%20-%20Learning%20-%20Tom%20Mitchell.pdf>>. Acesso em: 1/01/2023.

Russell, Stuart Jonathan, Inteligência artificial / Stuart Russell, Peter Norvig; tradução Regina Célia Simille. – Rio de Janeiro: Elsevier, 2010. Disponível em: < <https://www.cin.ufpe.br/~gtsa/Periodo/PDF/4P/SI.pdf>>. Acesso em: 1/01/2023.