

Todo list

validar o impacto da evasão no orçamento das IFES	3
estimar o prejuízo econômico da evasão para a UFC	3
adicionar exemplos de aplicação	7
é necessário traduzir credit assignment?	7
discutir	8
qual a escolhida no Mitchell? Lembrar de apresentar o exemplo	8
ponto muito confuso do capítulo! qual o valor inicial dos coeficientes? random? qual o valor de V_{treino} para o último estado do tabuleiro? usa um valor arbitrário? qual? acho que +100 win -100 loss e intuitivamente parece mesmo que a tendência é convergir	10
continuar a discussão dessa parte	10
definir o que quer dizer por modelo	11
melhorar	11

1 Introdução

O fenômeno evasão de discente consiste na interrupção de um processo de aprendizado de um discente antes de sua conclusão. Por exemplo, um discente que abandonou o curso de Computação, na UFC, no qual estava matriculado havia dois anos, pois precisou trabalhar para sustentar sua família e os horários das disciplinas eram incompatíveis com os horários do trabalho. Deste exemplo pode-se observar alguns dos atributos do fenômeno: o agente que interrompeu o processo, o discente, o curso, a instituição de ensino superior(IES), o tempo cursado e o motivo.

Sob a perspectiva de que o processo de aprendizado é um investimento e que o resultado esperado é a sua conclusão, o fenômeno evasão de discente pode ser considerado um problema: para a sociedade, com a frustração da expectativa de formação de profissionais e pesquisadores qualificados; para a instituição de ensino, caso tenha realizado investimentos em infraestrutura e em recursos humanos para atender a uma quantidade esperada de discentes ativos maior que a real, ocorrendo desperdício de recursos; caso seu orçamento seja ou função da quantidade de discentes ativos, no caso das instituições de ensino particulares, ou função da quantidade de discentes diplomados, no caso das instituições de ensino superior públicas; para o indivíduo que investiu tempo, dinheiro e dedicação, mas não terá os benefícios da conclusão da graduação, crítica no caso das profissões que exigem, para serem exercidas, diploma de graduação.

Em (PRESTES; FIALHO,) foi estimado um prejuízo econômico, decorrente da evasão de discentes de graduação no período de 2007 a 2012, para a UFPB, de R\$ 415.032.704,52. A estimativa considera perdidos os recursos financeiros investidos para manutenção do discente que não concluiu a graduação, utilizando a fórmula:

$$Perda\ Anual = n_evadidos \times t_permanencia \times v_aluno$$

onde:

- $n_evadidos$ representa a quantidade, por ano, média de discentes que evadiram no período, considerada a média aritmética do total de discentes que evadiram no período pela quantidade de anos do período.
- $t_permanencia$ representa o tempo de permanência esperado de um discente antes de evadir.
- v_aluno representa o custo corrente com hospital universitário por aluno corrente, indicador definido pelo TCU(TCU, 2009).

Aplicando a fórmula para os dados da UFC, obtemos o resultado

estimar o prejuízo econômico da evasão para a UFC

O estudo da evasão de discentes é motivado não apenas pelos problemas que dela podem decorrer mas também por diretrizes dos diversos níveis administrativos envolvidos com o processo.

validar
o
impacto
da
evasão
no
orçamento
das
IFES

No nível federal, redução da ocorrência desse fenômeno faz parte de uma das diretrizes do Programa de Apoio a Planos de Reestruturação e Expansão das Universidades Federais (REUNI), instituído pelo Decreto nº 6.096, de 24 de abril de 2007 (REPÚBLICA, 2007):

I - redução das taxas de evasão, ocupação de vagas ociosas e aumento de vagas de ingresso, especialmente no período noturno;

Na UFC, o instrumento de planejamento Plano de Desenvolvimento Institucional (UFC, 2013), para o período de 2013 a 2017, apresenta como um dos objetivos da política de assistência estudantil a redução da evasão; o programa de gestão da chapa eleita para reitoria no período de 2015 a 2019 (CUSTÓDIO, 2015) apresenta um conjunto de propostas que possuem como um dos objetivos a redução dos índices de evasão; o planejamento estratégico do Centro de Tecnologia da UFC para o período de 2015 a 2025 propõe a criação de uma equipe de apoio pedagógico para atuar no combate a problemas relacionados à evasão de discentes.

Em (LOBO, 2012) são apresentadas sete ações que ajudam a diminuir a ocorrência de evasão de discentes:

1. Estabelecer um grupo de trabalho encarregado de reduzir a evasão
2. Avaliar as estatísticas da evasão
3. Determinar as causas da evasão
4. Estimular a visão da IES centrada no aluno
5. Criar condições que atendam aos objetivos que atraíram os alunos
6. Tornar o ambiente e o trânsito na IES agradáveis aos alunos
7. Criar programa de aconselhamento e orientação dos aluno

Estas ações podem ser beneficiadas pela utilização de ferramentas de Aprendizado de Máquina, subárea de Inteligência Artificial, que estuda o desenvolvimento de programas cujas performances melhorem a partir de dados.

Para diminuir as taxas de evasão, uma das estratégias adotadas é a identificação precoce de discentes com grande tendência para abandonarem seus cursos e a execução de ações que minimizem tal tendência. A identificação pode ser conduzida por observação do comportamento e resultados dos discentes, de forma subjetiva, pelos docentes e coordenadores de cursos, por exemplo. Em estudo realizado no departamento de engenharia elétrica da Eindhoven University of Technology (DEKKER; PECHENIZKIY; VLEESHOUWERS, 2009), é relatado que em dezembro os discentes desse departamento recebem um aviso informando se são ou não aconselhados a continuarem no curso. Esse aviso é baseado na performance do discente no curso e em informações obtidas de professores do primeiro semestre e de discentes monitores. É relatado que o aviso parece ter bastante acurácia: geralmente discentes aconselhados a continuarem têm sucesso no próximo ano do curso, enquanto aqueles desaconselhados geralmente não continuam no curso. Dois problemas decorrem dessa forma de identificação: sendo conduzida por pessoas, essa forma de

identificação é limitada pelo conjunto de observações as quais o observador tem acesso; sendo subjetiva, seus resultados podem sofrer resistência para serem aceitos. A utilização de técnicas de aprendizado de máquina como forma de identificação pode contornar esses problemas, por, primeiro, fazer uso de dados registrados por sistemas de informação, provavelmente contendo informações mais amplas que as que uma pessoa pode observar; segundo, por fazer maior uso de dados registrados, sendo aceita mais facilmente como identificação objetiva. Nesse estudo foram utilizados diversos algoritmos de aprendizado de máquina com o objetivo de tentar detectar que um estudante irá abandonar seu curso. Foram utilizadas informações de discente referentes tanto ao período anterior ao seu ingresso na universidade, quanto ao posterior.

O presente trabalho objetiva avaliar a aplicabilidade de técnicas de aprendizado de máquina ao problema de evasão de discentes na UFC.

2 Aprendizado de máquina

2.1 Definição

Aprendizado de máquina é uma subárea de Inteligência Artificial que agrupa conhecimentos sobre algoritmos e técnicas que permitam que um programa melhore sua performance a partir de dados. Mais formalmente:

Um programa aprende a partir de uma experiência E , com relação a uma classe de tarefas T e a uma medida de performance P , se sua performance em tarefas da classe T , medida por P , melhora com a experiência E . (MITCHELL, 1997, p.2, tradução nossa)

adicionar exemplos de aplicação

2.2 Modelagem

Para exemplificar a modelagem de um programa com uma abordagem de aprendizado de máquina, (MITCHELL, 1997) apresenta uma sequência de passos para desenvolver um programa que aprenda a jogar xadrez, a ser utilizado para disputar um campeonato mundial de xadrez, tendo sua performance medida pela frequência de partidas que vencer.

Escolha da experiência

O primeiro passo consiste na escolha da experiência a partir da qual o programa irá aprender, denominada experiência de treinamento. (MITCHELL, 1997) apresenta três atributos da experiência de treinamento que devem ser levados em consideração por impactarem no sucesso do aprendizado: feedback, nível de controle sobre os exemplos de treinamento e quão bem a distribuição dos exemplos de treinamento representam a distribuição dos exemplos sobre os quais a performance final do programa será mensurada.

O atributo feedback representa quão direta é a informação fornecida pela experiência para o problema em questão, podendo assumir os valores feedback direto e feedback indireto. Por exemplo, a tupla estado do tabuleiro e o melhor movimento possível a partir desse estado é classificada como experiência de feedback direto: o programa irá atuar realizando movimentos e esse tipo de experiência informa diretamente qual o melhor movimento a ser executado. Já a tupla sequência de movimentos de uma partida e seu resultado final é classificada como experiência de feedback indireto: o resultado final da partida não fornece informação direta sobre a influência dos movimentos que nela foram executados. A atividade de determinar o grau de influência que elementos de uma experiência de feedback indireto têm sobre o resultado é denominada credit assignment.

Com relação ao nível de controle sobre os exemplos de treinamento, a experiência pode ser classificada como experiência selecionada por especialista, experiência sugerida pelo programa e analisada por um especialista e experiência selecionada e analisada pelo programa. Por exemplo, a experiência será do tipo selecionada por especialista se foi

é necessário traduzir credit assignment

selecionada por um jogador experiente de xadrez, que selecionou estados de tabuleiro e indicou que melhores movimentos poderiam ser feitos, a partir desses estados; será do tipo sugerida pelo programa e analisada por um especialista se o próprio programa selecionar estados de tabuleiro para serem analisadas por um jogador experiente de xadrez; será do tipo selecionada e analisada pelo programa se o programa utilizar o resultado de partidas que disputar consigo mesmo.

Com relação a quão bem a a distribuição dos exemplos de treinamento representam a distribuição dos exemplos sobre os quais a performance final do programa será mensurada, a experiência pode ser classificado como representativa, se a distribuição de seus exemplos representar a distribuição dos exemplos com os quais o programa efetivamente será utilizado, e não representativo, caso contrário. Por exemplo, a experiência é não representativa caso esteja limitada ao conjunto de partidas de apenas um jogador: considerando que o programa será utilizado em um campeonato mundial, do qual participam jogadores diversos, com estilos de jogo diversos, é esperado que o programa, treinado com essa experiência, depare com estados de tabuleiro que não encontrou no treinamento. (MITCHELL, 1997) ressalta que muito da teoria de aprendizado de máquina depende da assunção de que a experiência utilizada no treinamento é representativa.

Para o problema em análise, (MITCHELL, 1997) escolhe como experiência de treinamento um conjunto de partidas jogadas pelo programa contra ele mesmo.

Escolha da função alvo

O próximo passo é a escolha do tipo de conhecimento que deverá ser aprendido, representado por uma função denominada função alvo, e como ele será utilizado pelo programa. Considerando que o programa irá atuar como um jogador de xadrez, uma possível função a ser considerada é uma cujo domínio seja o conjunto de estados de tabuleiro e que retorne o melhor movimento a partir do estado de tabuleiro informado. Esse tipo de conhecimento depende da assunção de que, dado um estado de tabuleiro, existe um melhor movimento a ser executado. O problema de aprendizado dessa função depende, portanto, do problema de determinar quão um movimento influencia no resultado final de uma partida.

discutir

qual a escolhida no Mitchell? Lembrar de apresentar o exemplo

Outro exemplo de função alvo, a ser utilizada no exemplo, é uma que tenha como domínio o conjunto de estados de tabuleiro e retorne um número real, indicando quão bom o estado de tabuleiro informado é. O programa irá jogar verificando qual estado de tabuleiro maximiza o valor da função, considerando o conjunto de estados de tabuleiro que podem ser alcançados a partir do estado atual do tabuleiro e de todas jogadas válidas.

Escolha de uma representação para a função alvo

Após a escolha do tipo de conhecimento que deverá ser aprendido, é necessário definir como esse conhecimento será representado. Por exemplo, a função que associa um estado de tabuleiro a um número real pode assumir diversas formas: pode ser uma matriz contendo uma célula com um número real para cada estado de tabuleiro possível; um conjunto de regras que associe atributos do estado do tabuleiro a números reais;

uma função polinomial de atributos do estado do tabuleiro em números reais etc. Para dar continuidade ao detalhamento dos passos, (MITCHELL, 1997) escolhe a seguinte representação de função, denominada V :

$$V(t) = w_0 + w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + w_5x_5 + w_6x_6$$

Os coeficientes w_0 a w_6 são parâmetros do programa, a serem ajustados no aprendizado. As variáveis x_1 a x_6 possuem as seguintes definições:

- x_1 : número de peças pretas no tabuleiro
- x_2 : número de peças brancas no tabuleiro
- x_3 : número de reis pretos no tabuleiro
- x_4 : número de reis brancos no tabuleiro
- x_5 : número de peças pretas ameaçadas por peças brancas no tabuleiro
- x_6 : número de peças brancas ameaçadas por peças pretas no tabuleiro

Resumindo os passos até aqui realizados, temos:

- Tarefa: jogar xadrez
- Medida de performance: frequência de partidas do campeonato mundial de xadrez ganhas
- Experiência de treinamento: partidas disputadas pelo programa contra ele mesmo
- Função alvo: $V : \text{Estados De Tabuleiro} \rightarrow \mathbb{R}$
- Representação da função alvo: $\bar{V}(t) = w_0 + w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + w_5x_5 + w_6x_6$

Escolha de um algoritmo de aproximação

O próximo passo consiste na escolha de um algoritmo que irá realizar o treinamento da função, isto é, que, a partir da experiência de treinamento, irá ajustar seus coeficientes a fim de aproximá-la da função alvo. Para tanto, é necessário um conjunto de exemplos de treinamento composto por tuplas, $\langle b, V_{treino}(b) \rangle$, de estado de tabuleiro e o valor da função alvo para esse estado, respectivamente, b e $V_{treino}(b)$. Por exemplo, o seguinte exemplo representa um estado de tabuleiro onde o jogador com peças pretas ganhou:

$$\langle \langle x_1 = 3, x_2 = 0, x_3 = 1, x_4 = 0, x_5 = 0, x_6 = 0 \rangle, +100 \rangle$$

Para o desenvolvimento dos exemplos de treinamento, é necessário estimar um valor para cada estado de tabuleiro alcançado nas partidas disputadas pelo programa contra ele mesmo (experiência de treinamento escolhida), indicando quão bom aquele estado de tabuleiro é. A abordagem adotada no exemplo consiste em atribuir a $V_{treino}(b)$, caso seja um estado final, um valor arbitrário, caso seja um estado intermediário, o valor da função estimada, $\bar{V}(t)$, aplicado ao próximo estado do tabuleiro para o jogador ativo em b , $Sucessor(b)$:

$$V_{treino}(b) \leftarrow \bar{V}(t)(Sucessor(b)) \quad (2.1)$$

(MITCHELL, 1997) indica que, apesar de parecer estranho atribuir aos exemplos de treinamento valores estimados pela função que está sendo treinada, sob certas condições, pode ser provado que essa abordagem converge para uma estimativa perfeita de V_{treino} .

ponto muito confuso do capítulo! qual o valor inicial dos coeficientes? random? qual o valor de V_{treino} para o último estado do tabuleiro? usa um valor arbitrário? qual? acho que +100 win -100 loss e intuitivamente parece mesmo que a tendência é convergir

Para estimar a função alvo a partir dos dados de treinamento é necessário estimar seus pesos, w_0 a w_6 . Para tanto, primeiro é necessário definir uma medida de quão bem a função estimada se ajusta aos exemplos de treinamento. Uma medida comum é o erro quadrático, assim definido:

$$E = \sum_{(b, V_{treino}(b)) \in \text{dados de treinamento}} (V_{treino}(b) - V(b))^2$$

O problema de estimar a função alvo pode ser modelado então como o problema de encontrar os pesos w_0 a w_6 que minimizem o erro quadrático sobre os dados de treinamento. Um dos algoritmos que incrementalmente ajusta os pesos aos dados de treinamento, minimizando o erro quadrático, é o Least Mean Square. Esse algoritmo funciona atualizando iterativamente, para cada exemplo de treinamento, os pesos da função, ajustando-os com um valor proporcional a $(V_{treino}(b) - V(b))$: se o valor da função aplicado ao exemplo de treinamento, $V(b)$, for igual ao valor do exemplo de treinamento, $V_{treino}(b)$, o ajuste será nulo; caso seja maior, o ajuste será negativo, fazendo com que, para o dado de treinamento em questão, o valor da função diminua; caso seja menor, o incremento será positivo, fazendo com que, para o dado de treinamento em questão, o valor da função aumente.

continuar a discussão dessa parte

Conclusão

Nessa sequência de passos podem ser identificados quatro elementos de um programa que aprende:

- Sistema de performance: elemento responsável pela utilização do conhecimento aprendido para resolver uma tarefa. No exemplo, será responsável por determinar qual a próxima jogada, dado um estado de tabuleiro, utilizando a função que foi aprendida.
- Crítico: elemento responsável por receber como entrada um dado de treinamento e informar a que valor deve ser associado. No exemplo, o crítico é representado pela equação 2.1.
- Generalizador: elemento responsável por receber como entrada um conjunto de dados de treinamento e retornar uma função estimativa de uma função alvo. No exemplo, o generalizador é o Método dos Mínimos Quadrados.
- Representação da função alvo: elemento que define a estrutura da função que será utilizada como estimativa da função alvo. No exemplo, foi utilizada como representação uma combinação linear.

Outras configurações para esses elementos foram desenvolvidas. Por exemplo, como representação da função alvo pode-se utilizar um grafo em estrutura de árvore, denominado árvore de decisão. Cada nó seu que não seja folha possui uma regra que associa um dado a um de seus nós filhos. Os nós folhas são associados a um valor. Seu funcionamento consiste em apresentar um dado à regra de um nó, inicialmente o nó raiz, e recursivamente aplicar esse procedimento ao nó filho ao qual a regra associa o dado, até que seja alcançado um nó folha, cujo resultado associado é então retornado como o valor da função.

2.3 Modelos

definir o que quer dizer por modelo

representation + evaluation + optimization ([DOMINGOS, 2012](#))

2.4 Aplicações

melhorar

De acordo com (??), o conhecimento sobre aprendizado de máquina pode ser aplicado a diversas áreas, como, por exemplo, a reconhecimento de voz; a visão computacional, sendo utilizado no desenvolvimento de sistemas de reconhecimento facial; a controle de robôs.

Referências

- CUSTÓDIO, H. e. *Programa de gestão 2015-2019*. 2015. <<http://www.henry-custodio.com/programa.pdf>>.
- DEKKER, G. W.; PECHENIZKIY, M.; VLEESHOUWERS, J. M. Predicting students drop out: A case study. *International Working Group on Educational Data Mining*, ERIC, 2009.
- DOMINGOS, P. A few useful things to know about machine learning. *Communications of the ACM*, ACM, v. 55, n. 10, p. 78-87, 2012.
- LOBO, M. B. d. C. M. Panorama da evasão no ensino superior brasileiro: aspectos gerais das causas e soluções. *Associação Brasileira de Mantenedoras de Ensino Superior. Cadernos*, n. 25, 2012.
- MITCHELL, T. M. *Machine Learning*. 1. ed. New York, NY, USA: McGraw-Hill, Inc., 1997. ISBN 0070428077, 9780070428072.
- PRESTES, E. M. D. T.; FIALHO, M. G. D. A evasão no ensino superior globalizado e suas repercussões na gestão universitária.
- REPÚBLICA, P. da. *Decreto nº 6096*. 2007. <http://legislacao.planalto.gov.br/legisla/legislacao.nsf/Viw_Identificacao/DEC%206.096-2007?OpenDocument>.
- TCU. *Orientações para o cálculo dos indicadores de gestão - TCU*. 2009. <<http://portal.mec.gov.br/sesu/arquivos/pdf/indicadores.pdf>>.
- UFC. *Plano de Desenvolvimento Institucional*. 2013. <<http://ufc.br/a-universidade/documentos-oficiais/313-plano-de-desenvolvimento-institucional-pdi>>.