

FERRAMENTA ESTRATÉGICA PARA PREVISÃO DA EVASÃO DE ALUNOS NA FACULDADE TECNOLÓGICA DE PRESIDENTE PRUDENTE

Laisla Mirele Pulido da Silva Francisco¹

Daniel dos Santos Viais Neto²

Vanessa dos Anjos Borges³

Aluna do Curso Superior de Tecnologia em Análise e Desenvolvimento de Sistemas da FATEC

Presidente Prudente; e-mail: laisla.francisco@fatec.sp.gov.br¹

Professor da FATEC Presidente Prudente; e-mail: dv.netto@fatec.sp.gov.br²

Professor da FATEC Presidente Prudente; e-mail: vanesa.borges2@fatec.sp.gov.br³

Área do Conhecimento: Matemática aplicada e computacional

Palavras-chave: Aprendizado de máquina; Lagrange; Tomada de Decisão.

INTRODUÇÃO

A coleta de dados faz parte da natureza existencial do ser humano. Na era das cavernas, por exemplo, todas as decisões eram tomadas a partir do momento em que situações eram observadas ou ações eram executadas por partes, até que o melhor resultado pudesse ser alcançado. Com a evolução do ser humano e, consequentemente da tecnologia e da comunicação, todo dado produzido se torna crucial para que as organizações de qualquer natureza saibam mais sobre si mesmas.

Sendo um grande ativo das organizações, os dados podem prover muito mais do que fatos do cotidiano: podem fornecer previsões de futuro, a partir da utilização de ferramentas e técnicas que conseguem prever e simular situações futuras (VALENTIM, 2002). O Aprendizado de Máquina (*Machine Learning*) é uma dessas técnicas que podem ser utilizadas para extração de conhecimento a partir de uma coleção de dados.

Segundo Lopes e Costa (2017), o papel desempenhado por recursos computacionais para realização de cálculos como equações polinomiais são de grande auxílio, pois realizar o processo manualmente torna a prática inviável.

OBJETIVOS

O presente trabalho tem por objetivo apresentar um algoritmo desenvolvido com base em cálculos de Interpolação de Lagrange para realização de previsão de dados futuros da curva de evasões de alunos dos cursos da Faculdade de Tecnologia de Presidente Prudente com base nos dados pré-existentes, buscando auxiliar no processo de tomada de decisão da equipe de gestão escolar.

METODOLOGIA

Este trabalho caracteriza-se como uma pesquisa aplicada. Segundo Fleury e Werlang (2018, p.3), tal pesquisa concentra-se em torno dos problemas presentes nas atividades das instituições, organizações, grupos ou atores sociais e, está empenhada na elaboração de diagnósticos, identificação de problemas e busca de soluções.

Segundo Bassanezi (2011) a modelagem matemática é um processo que visa validar modelos matemáticos, com a abstração e generalização, conseguindo assim encontrar previsões tendenciosas. Tomados pela necessidade de levar as organizações a conseguirem os melhores resultados, a modelagem matemática será aplicada para obter conhecimento e solucionar problemas nesse projeto.

Nesse sentido, após a realização de pesquisas exploratórias nas principais bases de trabalhos científicos das áreas da Ciência da Computação e Matemática, assim como nos anais dos principais congressos relacionados à temática do projeto, sobre as temáticas *Machine Learning* e Interpolação de Lagrange,

foi desenvolvido um algoritmo na linguagem Python para validação da proposta de previsão dados. Para tal propósito, foram utilizados os dados informados pelo Ministério da Saúde sobre a pandemia de COVID19 (FRANCISCO; BORGES; NETO, 2020).

A partir da validação do algoritmo, foi desenvolvido um software utilizando a linguagem de programação C# com o .NET Framework versão 4.7.2, seguindo o padrão arquitetural MVC 5, para persistência de dados o *framework Entity* e sistema gerenciador de banco de dados SQLServer. Para a criação dos gráficos foi utilizada a biblioteca Google Charts.

RESULTADOS E DISCUSSÃO

O sistema desenvolvido ainda não foi hospedado em um servidor na nuvem, sendo disponibilizado apenas para acesso local. Ao acessar o sistema, são apresentadas opções que trazem informações sobre o sistema, como ele funciona e uma interface para acessá-lo. A proposta fundamental do sistema é que os gestores (coordenação e direção) da FATEC de Presidente Prudente possam ter uma ferramenta estratégica que forneça uma visão de futuro sobre as evasões, e que assim eles possam analisar estratégias para evitar que tal cenário se concretize.

Para isso, a cada novo trancamento realizado, um usuário do sistema irá lançar um novo registro, informando a qual curso o aluno pertence, qual o motivo levou ao trancamento e em qual ano e semestre esse trancamento foi realizado.

Para o gerenciamento dos dados no sistema, foram desenvolvidas três funções básicas, destinadas para inserção, alteração, consulta e exclusão (acrônimo em inglês popularmente conhecido como CRUD) dos dados de um trancamento. A Figura 2 apresenta essas interfaces.

Figura 1: Interfaces de gerenciamento de dados de Cursos, Motivos e Trancamentos

The figure consists of three vertically stacked screenshots of a web application interface. All three screenshots share a common header bar with the following navigation items: PREDICTION, Olá Administrador!, Dashboard, Cursos, Motivos, Trancamentos, and Log off. The first screenshot shows a form for creating a new course, with fields for NomeCurso (NomeCurso) and SiglaCurso (SiglaCurso). It also displays existing course information: Análise e Desenvolvimento de Sistemas (ADS) with Editar | Detalhes | Excluir links. The second screenshot shows a form for creating a new reason, with fields for Descricao (Descrição) and COVID (COVID). It also displays existing reason information: COVID-19 with Editar | Detalhes | Excluir links. The third screenshot shows a table for managing suspensions, with columns: NomeCurso (NomeCurso), Descricao (Descrição), Ano (Ano), and Semestre (Semestre). The table lists three entries for the course 'Análise e Desenvolvimento de Sistemas': COVID-19 (2020, 1º Semestre), COVID-19 (2020, 1º Semestre), and COVID-19 (2020, 2º Semestre). The last two rows are highlighted with a red border.

NomeCurso	Descricao	Ano	Semestre	
Análise e Desenvolvimento de Sistemas	COVID-19	2020	1º Semestre	Editar Detalhes Excluir
Análise e Desenvolvimento de Sistemas	COVID-19	2020	1º Semestre	Editar Detalhes Excluir
Análise e Desenvolvimento de Sistemas	COVID-19	2020	2º Semestre	Editar Detalhes Excluir

Fonte: Elaborado pelos autores

Para contextualizar o realce em vermelho apresentado na Figura 2, foram lançados três trancamentos realizados no curso de Análise e Desenvolvimento de Sistemas, dois no 1º Semestre de 2020 e um no 2º Semestre de 2020.

A partir dos lançamentos dos trancamentos, o algoritmo desenvolvido utilizando a Interpolação de Lagrange realiza a previsão da evasão para o próximo semestre que ainda não foi registrado no sistema. Para essa previsão, o usuário seleciona a opção Dashboard no sistema e depois os dados de qual curso deseja analisar. Assim, o sistema irá informar o agrupamento dos dados inseridos no sistema, por ano e semestre, e a respectiva quantidade de trancamentos realizados.

Figura 2: Interface de apresentação do resultado de previsão



Fonte: Elaborado pelos autores

Vale ressaltar que as previsões realizadas pelo algoritmo não levam em consideração os motivos que levaram as evasões, apesar de esses terem sido catalogados apenas para fins de observação. Cabe aos usuários realizar essa análise a partir do gráfico sobre os motivos dos trancamentos apresentado pela ferramenta. Os dados inseridos e apresentados nesse trabalho são fictícios, não apresentando a realidade dos trancamentos da unidade. A série de dados utilizada foi aplicada para a validação do algoritmo, pois, a previsão correta a partir da série de dados 2, 4 e 6 é o valor 8, demonstrando o funcionamento correto do algoritmo.

CONCLUSÕES

O uso da modelagem matemática neste trabalho demonstra que a aplicação de técnicas como o Polinômio de Lagrange para análise preditiva pode ser uma ferramenta importante da ciência de dados para o processo de tomada de decisões, pois fornece previsões aproximadas de realidades futuras, que podem ser analisados pelos gestores e os impactos consequentes dessas previsões sejam gerenciados. Para trabalhos futuros pretende-se analisar a implementação de cálculos e ferramentas que minimizem o erro entre dados reais e previsão para assim gerar um resultado mais preciso.

REFERÊNCIAS BIBLIOGRÁFICAS

- BASSANEZI, Betina Silvia Beozzo et al. Construção de um modelo matemático fuzzy para predizer o risco de vômitos pós-operatórios numa população pediátrica oncológica a partir da determinação dos fatores de risco. Tese de Doutorado. UNICAMP, 2011.
- CORTES, Corinna; VAPNIK, Vladimir. Support-vector networks. **Machine Learning**, v. 20, n. 3, p. 273-297, 1995.
- FLEURY, Maria Tereza Leme; WERLANG, Sérgio. **Pesquisa aplicada: reflexões sobre conceitos e abordagens metodológicas**. 2018.
- FRANCISCO, L. M. P. S.; BORGES, V. A.; NETO, D. S. V. APLICAÇÃO DE TÉCNICA DE MACHINE LEARNING PARA PREVISÃO DA CURVA DE DADOS DA COVID-19. In: **IX JORNACITEC-Jornada Científica e Tecnológica**. 2020.
- LOPES, P. Álvaro; COSTA, S. J Manuel. Comparação entre métodos de aproximação numérica utilizando o programa Matlab. **MARGENS – Revista Interdisciplinar**, versão digital, 2017.
- MENEZES, Isabella Barbosa; DUARTE, Marco Aparecido Queiroz. Métodos de interpolação polinomial. In: ANAIS DO ENCONTRO DE INICIAÇÃO CIENTÍFICA (ENIC), 8., 2010, Mato Grosso do Sul. **Anais eletrônicos...** Mato Grosso do Sul: UEMS, v. 1, n. 1(2), 2010.
- SILVA, Marcelo Mourão. **Uma abordagem evolucionária para aprendizado semisupervisionado em máquinas de vetores de suporte**. 2008, Dissertação (Mestrado Engenharia Elétrica). PPGEE/UFMG. Belo Horizonte, 2008.
- SPERANDIO, Décio; MENDES, João Teixeira; SILVA, Luiz Henry Monken. Cálculo Numérico – Características Matemáticas e Computacionais dos Métodos **Numéricos**, Prentice Hall, São Paulo, 2003.
- VALENTIM, Marta Lígia Pomim et al. Inteligência competitiva em organizações: dado, informação e conhecimento. **DataGramZero**, Rio de Janeiro, v. 3, n. 4, p. 1-13, 2002.