

Plano de Trabalho de Conclusão de Curso

Detecção de padrões em relação à presença dos estudantes no ensino básico brasileiro

UDESC – Centro de Ciências Tecnológicas
Departamento de Ciência da Computação
Bacharelado em Ciência da Computação
Turma 2023/1 – Joinville/SC

Daniella Martins Vasconcellos – `daniella.vasconcellos@edu.udesc.br`
Isabela Gasparini – `isabela.gasparini@udesc.br` (*orientadora*)
Elaine Harada Teixeira de Oliveira (UFAM) – `elaine@icomp.ufam.edu.br`
(*coorientadora*)

Março de 2023

Resumo

A educação deveria ser uma questão prioritária no Brasil e desde a criação do Ministério da Educação (MEC), políticas educacionais foram desenvolvidas para melhorar a qualidade do ensino no país. Para melhor gerenciamento de recursos e atendimento à população, é fundamental a coleta de informações e dados dos cidadãos que usufruem dos serviços. Esses dados coletados pelo MEC são utilizados por pesquisadores em todo o país para entender melhor o sistema educacional brasileiro e para desenvolver ou aprimorar políticas e programas que visam melhorar a qualidade da educação. Uma das várias formas de análise de dados é a ciência do reconhecimento de padrões, que tem como objetivo identificar padrões e regularidades em conjuntos de dados complexos, permitindo previsões e tomadas de decisão mais precisas e eficazes. Este trabalho tem como objetivo detectar padrões de estudantes do ensino básico brasileiro (que abrange a pré-escola, o ensino fundamental e o ensino médio) com base na presença em sala de aula. Para tal, será utilizado a metodologia de aprendizado de máquina de análise de séries temporais, que permitirá a identificação de padrões de frequência escolar dos estudantes ao longo do tempo. Como resultados esperados, espera-se que essa análise permita identificar possíveis fatores que contribuem para a evasão escolar e, assim, possibilitar a criação de políticas e programas mais eficazes para a promoção da permanência e do sucesso dos estudantes na escola.

Palavras-chave: *Reconhecimento de padrões, ensino básico, informática na educação.*

1 Introdução e Justificativa

Desde a criação do Ministério da Educação em 1930, o governo brasileiro desenvolveu políticas educacionais e programas que buscam melhorar a qualidade do ensino no país. Alguns dos programas incluem: o Programa Nacional de Alimentação Escolar (PNAE), que visa garantir a alimentação adequada dos estudantes de escolas públicas (BRASIL, 1955); o Programa Nacional do Livro Didático (PNLD), que visa fornecer livros didáticos gratuitos para estudantes de escolas públicas de todo o país e é considerado um dos maiores programas de distribuição gratuita de livros didáticos do mundo (BRASIL, 1985); o Plano Nacional de Educação (PNE), que é um conjunto de diretrizes e metas para o desenvolvimento da educação no país, estabelecendo objetivos e metas a serem alcançados em todas as etapas da educação (BRASIL, 2001); o Programa Universidade para Todos (ProUni), que visa democratizar o acesso ao ensino superior para estudantes de baixa renda, oferecendo bolsas de estudo integrais e parciais em universidades privadas para estudantes que atendam a certos critérios socioeconômicos (BRASIL, 2005); o Programa Mais Educação que visa ampliar a jornada escolar dos estudantes de escolas públicas e oferecer atividades complementares que estimulem o aprendizado e o desenvolvimento pessoal dos estudantes.

Como os programas governamentais do MEC têm como objetivo o fornecimento de serviços públicos essenciais para atender a variadas necessidades educacionais da população, é fundamental que se colete informações e dados sobre os cidadãos que usufruem dos serviços para melhor utilização de recursos e retorno à população. Por exemplo, o Prouni acessa dados socioeconômicos dos estudantes que estão em bases governamentais para poder analisar possíveis dispensas de apresentação de documentos (como comprovação de renda familiar mensal bruta ou a situação da pessoa com deficiência) (BRASIL, 2005).

A coleta de dados do MEC é utilizada por pesquisadores em todo o país para entender melhor o sistema educacional brasileiro, até para tecer críticas sobre seus programas (NASSAR; COUTO; JÚNIOR, 2021) (CARDOSO; BARRETTO; GIANEZINI, 2019) ou para compreender como os indicadores educacionais podem afetar outros programas governamentais (SANTOS et al., 2019). Os dados coletados pelo MEC incluem informações sobre a estrutura das escolas, o desempenho dos estudantes, a formação dos professores, o financiamento da educação e outros aspectos importantes da educação no Brasil. Esses dados são usados para realizar estudos e pesquisas que ajudam a identificar problemas e desafios, bem como para desenvolver políticas e programas que visam melhorar a qualidade da educação brasileira (CARREIRO, 2020; RIBEIRO; GUERRA, 2019). Além disso, os dados do MEC são fundamentais para a elaboração de índices de qualidade da educação, como o Índice de Desenvolvimento da Educação Básica (IDEB), que mede o desempenho dos estudantes em todo o país.

Os pesquisadores têm utilizado os dados do MEC para diversos fins, como por exemplo, para avaliar a eficácia de programas educacionais (VARGAS; ZUCCARELI, 2021), para identificar as necessidades de formação de professores, ou para analisar as desigualdades educacionais no país e desenvolver políticas públicas que buscam melhorar a qualidade da educação (SANTOS et al., 2019). Esses dados são essenciais para a elaboração de políticas e programas educacionais eficazes e para o monitoramento contínuo da qualidade da educação brasileira, permitindo aos pesquisadores,

gestores e educadores entenderem melhor a realidade educacional do país.

Uma das várias formas de análise de dados é a ciência do reconhecimento de padrões, que na computação é uma área interdisciplinar que envolve conhecimentos de matemática, estatística e inteligência artificial. Esse campo de estudo busca identificar padrões e regularidades em conjuntos de dados complexos, permitindo que sejam feitas previsões e tomadas de decisão mais precisas e eficazes (PAOLANTI; FRONTONI, 2020). O histórico da detecção de padrões na computação remonta à década de 1950 com estudos focados em padrões de imagens (SELFRIEDGE, 1955), e desde então muitos outros métodos foram conquistados, como a análise de séries temporais (GURALNIK; SRIVASTAVA, 1999), aprendizagem de máquina e mineração de dados.

Hoje em dia, a detecção de padrões tem se mostrado uma área de grande relevância para o desenvolvimento de tecnologias inovadoras, como carros autônomos (TSELENTIS; PAPADIMITRIOU, 2023), assistentes virtuais (SAYED; ASHOUR, 2021), reconhecimento facial (CHEN et al., 2021), sistemas de recomendação (AHMED; IMTIAZ; KHAN, 2018) e até para a medicina personalizada, em que a análise de grandes conjuntos de dados pode ajudar a identificar tratamentos mais eficazes e personalizados para cada paciente. Sendo assim, é uma das áreas mais promissoras da computação e tem contribuído significativamente para o avanço de diferenciadas tecnologias e diversos campos de estudo. No contexto da educação brasileira, o grande volume da coleta de dados do MEC pode permitir que padrões de comportamento e aprendizado sejam identificados, por exemplo, por meio da análise de dados de desempenho dos estudantes em diferentes disciplinas, que pode inferir padrões de dificuldades em determinadas matérias, indicando a necessidade de ajustes na metodologia de ensino ou na grade horária.

Sendo assim, é possível inferir que a detecção de padrões possa ser aplicada para a análise da frequência dos estudantes em sala de aula, e ainda de variados modos, como por reconhecimento facial (BUDIMAN et al., 2023) ou por dispositivos *wearables* (FERREIRA; RODRIGUEZ; MOTTI, 2020). A coleta de dados de presença dos estudantes pode ser utilizada para detectar padrões de ausências, indicando a necessidade de intervenções para melhorar a participação dos estudantes nas aulas, já que a baixa frequência de estudantes está relacionada com maiores índices de evasão escolar. O absentismo, por sua vez, tem distintos e complexos fatores de risco, incluindo vulnerabilidade social, abuso de substâncias, baixo envolvimento dos pais na escola, a falta de correspondência entre as abordagens educacionais e as necessidades dos estudantes, a carência de iniciativas governamentais efetivas e a exploração de trabalho infantil (NASCIMENTO et al., 2020; GUBBELS; PUT; ASSINK, 2019). Neste contexto, a análise dos dados referentes a presença escolar dos estudantes pode ser uma ferramenta útil, permitindo o aprimoramento de práticas pedagógicas com enfoque em retenção escolar. É importante destacar, no entanto, que a coleta de dados dos estudantes deve ser realizada de forma ética e responsável, respeitando a privacidade e os direitos dos estudantes.

Tendo em mente as problemáticas anteriormente apresentadas, o proposto trabalho de conclusão de curso tem como objetivo encontrar padrões de características de estudantes do ensino básico (o que abrange a pré-escola, o ensino fundamental e o ensino médio (BRASIL, 2013)) com base na sua presença em sala de aula. Pretende-se coletar informações dos estudantes (como notas, gênero,

idades, presenças em sala de aula) para identificar uma possível correlação com evasão escolar. O trabalho está inserido no projeto “Gestão Inteligente de Atores Educacionais da Escola com Dados Abertos Conectados”, no qual é coordenado pelo Núcleo de Excelência em Tecnologias Sociais da Universidade Federal de Alagoas (NEES-UFAL), com participação da UDESC e da UFAM.

2 Objetivos

Objetivo Geral: Encontrar padrões de características de estudantes do ensino básico brasileiro com base na presença em sala de aula.

Objetivos Específicos:

- Investigar sistemas de presença na literatura;
- Realizar levantamento bibliográfico sobre algoritmos de reconhecimento de padrões relacionados à mineração de dados;
- Estudar as bases de dados disponibilizadas pelo NEES sobre a presença dos estudantes do ensino básico brasileiro;
- Propor um modelo para a análise de dados sobre estudantes do ensino básico;
- Descobrir padrões existentes nas bases de dados.

3 Metodologia

Este trabalho de conclusão de curso possui natureza aplicada, e é caracterizado como um trabalho de caráter experimental, pois será necessária a aplicação de técnicas que serão analisadas no decorrer da pesquisa. A abordagem será na forma quantitativa, já que serão estudados dados estatísticos que posteriormente serão traduzidos em visualizações.

A pesquisa iniciará com um levantamento bibliográfico para investigar como é realizada a detecção e cobrança de presença de alunos no ensino básico brasileiro atualmente, assim como técnicas de reconhecimento de padrões dentro da área de aprendizado de máquina.

Após ter sido realizada a fundamentação teórica, será feita a análise dos dados obtidos, aplicando as técnicas encontradas de reconhecimento de padrões para encontrar modelos semelhantes de tipos de estudantes.

Para atingir os objetivos propostos, serão realizadas as seguintes etapas:

1. Estudar sistemas atuais de checagem de presença de estudantes em sala de aula;
2. Estudar diferentes algoritmos de reconhecimento de padrões para grandes bases de dados;
3. Fazer levantamento de trabalhos relacionados sobre detecção de presença estudantil em sala de aula;

4. Estudar as bases de dados dos estudantes do ensino básico disponibilizadas pelo NEES;
5. Definir algoritmo(s) de reconhecimento de padrões que seja(m) mais apropriado(s) para as bases de dados disponibilizadas pelo NEES;
6. Aplicar algoritmo(s) de reconhecimento de padrões;
7. Gerar visualizações;
8. Analisar os resultados;
9. Escrever o trabalho de conclusão de curso.

4 Cronograma proposto

Etapas	2023/1					2023/2				
	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
1										
2										
3										
4										
5										
6										
7										
8										
9										

5 Linha e Grupo de Pesquisa

Este trabalho pertence ao Grupo de Pesquisa em Informática na Educação (GPIE) e está relacionado ao projeto “Gestão Inteligente de Atores Educacionais da Escola com Dados Abertos Conectados” do NEES com parceria da UDESC e da UFAM.

6 Forma de Acompanhamento/Orientação

O acompanhamento das atividades desenvolvidas será realizada em reuniões semanais, presenciais e/ou via vídeo-chamada, com até uma hora de duração, tendo a possibilidade de agendar encontros extras, caso necessário. Também serão utilizados recursos como e-mail, para orientação ao longo da semana ou em dúvidas casuais.

O orientador solicitará tarefas para o orientando a serem cumpridas conforme o cronograma. Caso o aluno não cumpra alguma tarefa e/ou cumpra após o prazo estipulado (de forma que o orientador não possua tempo hábil para realizar as correções necessárias), o orientador poderá

solicitar seu desligamento da orientação. O desligamento também poderá ser solicitado em casos de não cumprimento das datas e horários de reunião por parte do aluno.

O aluno se compromete a participar das reuniões do grupo de pesquisa GPIE quando, eventualmente, será requerido a apresentar o andamento de seu trabalho aos demais. Todo material escrito pelo acadêmico deve ser submetido ao orientador com antecedência para ser avaliado e sofrer alterações e correções. Todo material (artigos, textos produzidos, código fontes) deverão estar disponíveis em forma editável em diretório compartilhado ao longo de todo o trabalho.

Referências

AHMED, M.; IMTIAZ, M. T.; KHAN, R. Movie recommendation system using clustering and pattern recognition network. In: **2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)**. [S.l.: s.n.], 2018. p. 143–147.

BRASIL. **Decreto nº 37.106, de 31 de Março de 1955**. 1955. Diário Oficial da União. <https://www2.camara.leg.br/legin/fed/decret/1950-1959/decreto-37106-31-marco-1955-332702-publicacaooriginal-1-pe.html>.

BRASIL. **Decreto nº 91.542, de 19 de Agosto de 1985**. 1985. Diário Oficial da União. <https://www2.camara.leg.br/legin/fed/decret/1980-1987/decreto-91542-19-agosto-1985-441959-publicacaooriginal-1-pe.html>.

BRASIL. **Lei Nº 010172, de 9 de janeiro de 2001**. 2001. Diário Oficial da União. https://www.planalto.gov.br/ccivil_03/leis/leis_2001/l10172.htm.

BRASIL. **Lei Nº 11.096, de 13 de janeiro de 2005**. 2005. Diário Oficial da União. https://www.planalto.gov.br/ccivil_03/_ato2004-2006/2005/lei/l11096.htm.

BRASIL. **Lei Nº 12.796, de 4 de abril de 2013**. 2013. Diário Oficial da União. https://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2013/Lei/L12796.htm#art1.

BUDIMAN, A. et al. Student attendance with face recognition (lbph or cnn): Systematic literature review. **Procedia Computer Science**, v. 216, p. 31–38, 2023. ISSN 1877-0509. 7th International Conference on Computer Science and Computational Intelligence 2022. Disponível em: <https://www.sciencedirect.com/science/article/pii/S187705092202186X>.

CARDOSO, M. A. C.; BARRETTO, L. A. M.; GIANEZINI, K. **Políticas públicas no século XXI**. [S.l.]: UNESCO, 2019. 288-298 p.

CARREIRO, F. B. Estratégias docentes em prol da melhoria da qualidade do ensino na reéde municipal de são luís/ma. **Reunião Científica Regional Nordeste da Associação Nacional de Pesquisa e Pós-Graduação em Educação**, 2020. ISSN 2595-7945.

CHEN, W. et al. Yolo-face: a real-time face detector. *Vis Comput*, n. 37, p. 805–813, 2021.

FERREIRA, P.; RODRIGUEZ, C.; MOTTI, V. Wearables para coleta de dados de estudantes em ambiente escolar: Mapeamento sistemático. In: **Anais do XXXI Simpósio Brasileiro de Informática na Educação**. Porto Alegre, RS, Brasil: SBC, 2020. p. 1353–1362. ISSN 0000-0000. Disponível em: <https://sol.sbc.org.br/index.php/sbie/article/view/12891>.

GUBBELS, J.; PUT, C. E. van der; ASSINK, M. Risk factors for school absenteeism and dropout: A meta-analytic review. **Journal of Youth and Adolescence**, v. 48, n. 48, 2019. ISSN 1637–1667. Disponível em: <https://doi.org/10.1007/s10964-019-01072-5>.

GURALNIK, V.; SRIVASTAVA, J. Event detection from time series data. In: . [S.l.: s.n.], 1999. p. 33–42.

NASCIMENTO, J. C. d. S. et al. Fracasso escolar e evasão no ensino médio no brasil: estado do conhecimento. **Revista Educar Mais**, v. 4, n. 2, p. 379–393, set. 2020. Disponível em: <https://periodicos.ifsul.edu.br/index.php/educarmais/article/view/1823>.

NASSAR, L. M.; COUTO, M. H. D. C.; JÚNIOR, G. A. P. Financiamento público (fies e prouni) para o ensino de medicina no brasil: Uma revisão da literatura e as distorções criadas. **Educação em Revista**, Faculdade de Educação da Universidade Federal de Minas Gerais, v. 37, n. Educ. rev., 2021 37, p. e25246, 2021. ISSN 0102-4698. Disponível em: <https://doi.org/10.1590/0102-469825246>.

PAOLANTI, M.; FRONTONI, E. Multidisciplinary pattern recognition applications: A review. **Computer Science Review**, v. 37, p. 100276, 2020. ISSN 1574-0137. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1574013719300899>.

RIBEIRO, W. L.; GUERRA, M. d. G. A. G. A. V. Avaliação de cursos a partir do SINAES: uma análise para melhoria da qualidade na Universidade Federal da Paraíba. **Revista Educação em Questão**, scielo, v. 57, 07 2019. ISSN 0102-7735. Disponível em: http://educa.fcc.org.br/scielo.php?script=sci_arttext&pid=S0102-77352019000300007&nrm=iso.

SANTOS, M. C. S. et al. Programa bolsa família e indicadores educacionais em crianças, adolescentes e escolas no brasil: revisão sistemática. **Ciência & Saúde Coletiva**, ABRASCO - Associação Brasileira de Saúde Coletiva, v. 24, n. Ciênc. saúde coletiva, 2019 24(6), p. 2233–2247, Jun 2019. ISSN 1413-8123. Disponível em: <https://doi.org/10.1590/1413-81232018246.19582017>.

SAYED, J.; ASHOUR, M. W. Digital intelligent virtual assistant (diva) with natural speech and accent recognition. **IET Conference Proceedings**, Institution of Engineering and Technology, p. 170–175(5), 2021. Disponível em: <https://digital-library.theiet.org/content/conferences/10.1049/icp.2022.0334>.

SELFIDGE, O. G. Pattern recognition and modern computers. In: **Proceedings of the March 1-3, 1955, Western Joint Computer Conference**. New York, NY, USA: Association for Computing Machinery, 1955. (AFIPS '55 (Western)), p. 91–93. ISBN 9781450378567. Disponível em: <https://doi.org/10.1145/1455292.1455310>.

TSELENTIS, D. I.; PAPADIMITRIOU, E. Driver profile and driving pattern recognition for road safety assessment: Main challenges and future directions. **IEEE Open Journal of Intelligent Transportation Systems**, v. 4, p. 83–100, 2023.

VARGAS, H. M.; ZUCCARELI, C. A nova face da docência: uma proposta de revisão do censo da educação superior. **Estudos em Avaliação Educacional**, scielo, v. 32, 00 2021. ISSN 0103-6831. Disponível em: http://educa.fcc.org.br/scielo.php?script=sci_arttext&pid=S0103-68312021000100103&nrm=iso.

Isabela Gasparini
Orientadora

Daniella Martins Vasconcellos

Isabela Gasparini
(Líder do grupo de pesquisa)