

Análises classificatórias sobre os fatores de risco de evasão escolar

Daniella Martins Vasconcellos

Isabela Gasparini
Orientadora

Elaine Harada Teixeira de Oliveira
(UFAM)
Coorientadora

30/11/2023
Trabalho de Conclusão de Curso 2



Roteiro da Apresentação

Introdução



Justificativa, objetivos e metodologia

Fundamentação Teórica



Educational Data Mining, Reconhecimento de Padrões, Ética

Trabalhos Relacionados



Apresentação de 3 trabalhos

Desenvolvimento



Análises da base, estatísticas, classificatórias

Resultados



Conclusões finais

Referências



Referências utilizadas na apresentação

1. Introdução

Contextualização



- O Ministério da Educação (MEC) coordena e executa várias políticas públicas com o objetivo de fornecer serviços públicos essenciais para a população na área da educação;
- Coleta de informações sobre os cidadãos para usufruírem dos serviços;
- Na ciência: Coleta de dados educacionais é usada para avaliar a qualidade do ensino e a efetividade de políticas públicas na área da educação em diferentes regiões do Brasil (SANTOS et al., 2019; MONTEIRO; FERREIRA; TEIXEIRA, 2009);



Contextualização

- Com a ciência de **reconhecimento de padrões** é possível fazer previsões e tomadas de decisões de forma mais estratégica;
- Reconhecer padrões por meio da frequência de estudantes pode permitir o **aprimoramento de práticas pedagógicas com enfoque em retenção escolar**.



TED 11476 NEES

- Projeto celebrado entre o Núcleo de Excelência em Tecnologias Sociais (NEES), coordenado pela Universidade Federal de Alagoas (UFAL), e o Ministério da Educação (MEC) por meio do Plano de Trabalho do Termo de Execução Descentralizada (TED) nº 11476/2022;
- Meta geral: desenvolvimento de uma solução de gestão inteligente para controle de frequência de alunos do ensino básico, com alto potencial de adesão das escolas, visando o **enfrentamento do abandono e da evasão escolar** e ampliar a atuação da solução automatizada para **melhorar a qualidade da educação básica no Brasil**;
- Início em 2022 e previsão de encerramento em 2024.



TED 10974 NEES

- Projeto celebrado entre o Núcleo de Excelência em Tecnologias Sociais (NEES), coordenado pela Universidade Federal de Alagoas (UFAL), e o Ministério da Educação (MEC) por meio do Plano de Trabalho do Termo de Execução Descentralizada (TED) nº 10974/2022;
- Meta geral: desenvolvimento de um sistema de alerta preventivo dos riscos de abandono e evasão escolar, capaz de **identificar precocemente os estudantes com maior risco de abandono e evasão escolar**, permitindo que sejam tomadas medidas preventivas para evitar que isso aconteça. O **Sistema de Alerta Preventivo (SAP)** será baseado em dados coletados sobre o desempenho dos alunos, como notas, frequência e comportamento.

Objetivo Geral

Descobrir os fatores de risco de evasão de estudantes com base em análises classificatórias de dados coletados.



Objetivos Específicos

- Realizar levantamento bibliográfico sobre algoritmos de reconhecimentos de padrões relacionados à mineração de dados;
- Estudar o Instrumento de Avaliação dos Fatores de Risco de Evasão Escolar (IAFREE) fornecido pelo NEES e a base gerada por ele;
- Realizar análises estatísticas sobre a base;
- Realizar análises classificatórias utilizando algoritmos de *machine learning* sobre a base.

2. Fundamentação Teórica

01 Educational Data Mining (EDM)

Área de pesquisa que analisa grandes conjuntos de dados educacionais, se relacionando com Learning Analytics (LA).

02 Reconhecimento de padrões

Estratégia que permite identificar regularidades, tendências ou estruturas subjacentes, e seus estudos de visualização de dados.

03 Ética

Discussões sobre valores éticos relacionados à coleta e análise de dados.

01

Educational Data Mining (EDM)

- Foca no desenvolvimento de métodos para explorar dados educacionais únicos e em grande escala, a fim de compreender melhor os alunos e os ambientes de aprendizagem (KOÇ, 2017);
- Relacionada à **Learning Analytics**, que coleta, analisa e interpreta dados de aprendizagem com ênfase na intervenção e participação humana.

01

Educational Data Mining (EDM)

- EDM e LA têm como objetivo a **análise de dados educacionais para entender o processo de ensino e aprendizagem e melhorá-lo** (MOISSA, GASPARINI, KEMCZKINSKI, 2015);
- Tomadores de decisão podem utilizar técnicas de **mineração de dados e identificação de padrões** para analisar comportamentos e preferências dos estudantes.

02

Reconhecimento de padrões

- Métodos computacionais que permitem **identificar regularidades ou padrões** em dados;
- Técnicas: Clustering, **Regras de Associação**, Time series, Sequential Pattern analysis;
- Algoritmos: Decision Tree, Random Forest, Naive Bayes Categórico, AdaBoost, K-Nearest-Neighbors;
- Métricas: Acurácia, Precisão, Recall Score, F1 Score.



3. Trabalhos relacionados

| Critério | Contexto do Estudo | Métodos Utilizados | Abordagem de Análise | Principais Resultados |
|----------------------------|--------------------------------|--|-----------------------------|-----------------------------------|
| Queiroga et al. (2020) | Curso técnico a distância | Algoritmo genético | Análise de interações | Otimização de hiperparâmetros |
| Lee e Chung (2019) | Universidade na Coreia do Sul | <i>Ensemble stacking</i> | Análise baseada em big data | Previsão de abandono escolar |
| Niyogisubizo et al. (2022) | Universidade em Nitra | <i>Ensemble stacking</i> | Validação cruzada | Alta precisão de previsão |
| Presente trabalho | Sistema educacional brasileiro | Random Forest, Decision Tree, Naive Bayes, AdaBoost, KNN | Análise classificatória | Identificação de fatores de risco |

Tabela 1. Comparação entre os estudos relacionados

4. Desenvolvimento

4.1 Contexto



- O estágio inicial da pesquisa foi estudar sistemas de presença segundo a TED 11476, porém houve a indisponibilidade de dados;
- Contato com o TED 10974, Sistema Altera Preventivo;
- Dados do **Instrumento de Avaliação dos Fatores de Risco de Evasão Escolar (IAFREE-R)**;
- Opção por realizar análises de *machine learning* com classificação, sem reconhecimento de padrões, mas utilizando os mesmos algoritmos e métricas estudados.

4. Desenvolvimento

4.2 Base Disponibilizada



- Questionário aplicado para 17.110 estudantes de todas as regiões do Brasil;
- 308 escolas diferentes;
- Questionário sociodemográfico;
- 36 perguntas relacionadas a possíveis motivos de evasão em escala de Likert variando de 1 a 7;
 - 1 (Discordo Totalmente) e 7 (Concordo Totalmente);
 - Sem resposta = 0.
- Perguntas divididas em grupos de **dimensões** que possuem subgrupos **fatores**.

4. Desenvolvimento

4.2 Base Disponibilizada

| | | |
|-----|---|--|
| Q1 | Às vezes, me sinto triste ou deprimido(a), e isso me faz pensar em deixar a escola. | |
| Q2 | Sinto que sou incapaz de concluir meus estudos e por isso penso em abandonar a escola. | |
| Q3 | A alimentação oferecida na escola me faz pensar em não ir para ela. | |
| Q4 | Pensei em me afastar da escola porque ela não tem equipamentos adequados para lidar com as condições do clima (ventilador, aquecedor, ar condicionado). | |
| Q5 | Já pensei em abandonar a escola porque não vejo como ela vai me ajudar a mudar minha vida para melhor. | |
| Q6 | Pensei em deixar a escola pois ela não trata dos assuntos atuais. | |
| Q7 | Já considerei abandonar a escola por ter tido e/ou ter um problema de saúde. | |
| Q8 | Já pensei em deixar a escola porque não tenho aulas sobre a história ou questões importantes da minha comunidade ou cidade. | |
| Q9 | Pensei em sair da escola porque não tenho um bom espaço em casa para me concentrar nos estudos. | |
| Q10 | Já pensei em abandonar a escola porque ela não me prepara para os empregos que quero ter no futuro. | |
| Q11 | Já pensei em deixar a escola porque ela não respeita a religião que eu pratico. | |
| Q12 | Pensei em abandonar a escola porque os(as) professores(as) faltam muito. | |
| Q13 | Pensei em me afastar da escola pela quantidade de regras que ela tem. | |
| Q14 | Já pensei em deixar a escola porque com frequência ela é alvo de violência (vandalismo, assaltos, pichações, toque de recolher e etc). | |
| Q15 | Já faltei à escola ou deixei de fazer lição de casa por ter que ajudar em atividades de casa (cozinhar, limpar, cuidar de irmãos etc). | |
| Q16 | Alguém da minha família já sugeriu que eu deixasse a escola. | |

| | | |
|-----|---|--|
| Q17 | Pensei em abandonar a escola porque não tenho dinheiro suficiente para vir, ou porque o caminho até aqui é muito difícil. | |
| Q18 | Penso em deixar a escola porque ela não me permite fazer atividades como música, dança, teatro, desenho. | |
| Q19 | Pensei em abandonar a escola porque as salas têm mais estudantes do que os(as) professores(as) conseguem dar atenção. | |
| Q20 | Pensei em abandonar a escola porque as aulas são repetitivas e cansativas. | |
| Q21 | Já pensei em abandonar a escola por ter tido e/ou ter um problema de saúde na família. | |
| Q22 | Pensei em abandonar a escola por não poder praticar os esportes que eu queria. | |
| Q23 | Não poder usar internet na escola é algo que me faz querer me afastar. | |
| Q24 | Já pensei em deixar a escola porque meus/minhas colegas não me tratam bem. | |
| Q25 | Falto às aulas mais do que é permitido. | |
| Q26 | Acredito que tenho mais dificuldade de acompanhar as aulas do que os outros colegas. | |
| Q27 | Minhas notas em Português e/ou Matemática estão abaixo da média. | |
| Q28 | Possuo o material escolar adequado para acompanhar as aulas (lápiz, caderno, etc.). | |
| Q29 | Meus/Minhas colegas frequentemente abusam de álcool e/ou drogas. | |
| Q30 | Meus pais/responsáveis não podem vir à escola quando são chamados | |
| Q31 | Já aconteceu de eu não poder continuar na escola por causa de uma gravidez, seja minha ou da minha parceira. | |
| Q32 | Minha mãe, ou a pessoa que cuida de mim, não terminou a escola. | |
| Q33 | Já tive/tenho contato com o Conselho Tutelar. | |
| Q34 | É comum que na minha escola estudantes enfrentem problemas com a justiça. | |
| Q35 | Não estou na série que deveria estar para a minha idade. | |
| Q36 | Não possuo o fardamento escolar adequado. | |

Imagem 1. Perguntas do IAFREE

4. Desenvolvimento

4.2 Base Disponibilizada


| | | |
|-----|---|--|
| Q1 | Às vezes, me sinto triste ou deprimido(a), e isso me faz pensar em deixar a escola. | |
| Q2 | Sinto que sou incapaz de concluir meus estudos e por isso penso em abandonar a escola. | |
| Q3 | A alimentação oferecida na escola me faz pensar em não ir para ela. | |
| Q4 | Pensei em me afastar da escola porque ela não tem equipamentos adequados para lidar com as condições do clima (ventilador, aquecedor, ar condicionado). | |
| Q5 | Já pensei em abandonar a escola porque não vejo como ela vai me ajudar a mudar minha vida para melhor. | |
| Q6 | Pensei em deixar a escola pois ela não trata dos assuntos atuais. | |
| Q7 | Já considerei abandonar a escola por ter tido e/ou ter um problema de saúde. | |
| Q8 | Já pensei em deixar a escola porque não tenho aulas sobre a história ou questões importantes da minha comunidade ou cidade. | |
| Q9 | Pensei em sair da escola porque não tenho um bom espaço em casa para me concentrar nos estudos. | |
| Q10 | Já pensei em abandonar a escola porque ela não me prepara para os empregos que quero ter no futuro. | |
| Q11 | Já pensei em deixar a escola porque ela não respeita a religião que eu pratico. | |
| Q12 | Pensei em abandonar a escola porque os(as) professores(as) faltam muito. | |
| Q13 | Pensei em me afastar da escola pela quantidade de regras que ela tem. | |
| Q14 | Já pensei em deixar a escola porque com frequência ela é alvo de violência (vandalismo, assaltos, pichações, toque de recolher e etc). | |
| Q15 | Já faltei à escola ou deixei de fazer lição de casa por ter que ajudar em atividades de casa (cozinhar, limpar, cuidar de irmãos etc). | |
| Q16 | Alguém da minha família já sugeriu que eu deixasse a escola. | |

| | | |
|-----|---|--|
| Q17 | Pensei em abandonar a escola porque não tenho dinheiro suficiente para vir, ou porque o caminho até aqui é muito difícil. | |
| Q18 | Penso em deixar a escola porque ela não me permite fazer atividades como música, dança, teatro, desenho. | |
| Q19 | Pensei em abandonar a escola porque as salas têm mais estudantes do que os(as) professores(as) conseguem dar atenção. | |
| Q20 | Pensei em abandonar a escola porque as aulas são repetitivas e cansativas. | |
| Q21 | Já pensei em abandonar a escola por ter tido e/ou ter um problema de saúde na família. | |
| Q22 | Pensei em abandonar a escola por não poder praticar os esportes que eu queria. | |
| Q23 | Não poder usar internet na escola é algo que me faz querer me afastar. | |
| Q24 | Já pensei em deixar a escola porque meus/minhas colegas não me tratam bem. | |
| Q25 | Falto às aulas mais do que é permitido. | |
| Q26 | Acredito que tenho mais dificuldade de acompanhar as aulas do que os outros colegas. | |
| Q27 | Minhas notas em Português e/ou Matemática estão abaixo da média. | |
| Q28 | Possuo o material escolar adequado para acompanhar as aulas (lápiz, caderno, etc.). | |
| Q29 | Meus/Minhas colegas frequentemente abusam de álcool e/ou drogas. | |
| Q30 | Meus pais/responsáveis não podem vir à escola quando são chamados | |
| Q31 | Já aconteceu de eu não poder continuar na escola por causa de uma gravidez, seja minha ou da minha parceira. | |
| Q32 | Minha mãe, ou a pessoa que cuida de mim, não terminou a escola. | |
| Q33 | Já tive/tenho contato com o Conselho Tutelar. | |
| Q34 | É comum que na minha escola estudantes enfrentem problemas com a justiça. | |
| Q35 | Não estou na série que deveria estar para a minha idade. | |
| Q36 | Não possuo o fardamento escolar adequado. | |

Imagem 1. Perguntas do IAFREE

4. Desenvolvimento

4.2 Base Disponibilizada



| Dimensão de Risco | Fatores de Risco | Cód. | Itens |
|--|--|------|--|
| Estudante – Família (E-FAM) | Suporte Familiar (E-FAM1) | Q16 | Alguém da minha família já sugeriu que eu deixasse a escola. |
| | | Q32 | Minha mãe, ou a pessoa que cuida de mim, não terminou a escola. |
| | | Q30 | Meus pais/responsáveis não podem vir à escola quando são chamados. |
| | Gravidez/parentalidade/ atividades de cuidado (E-FAM2) | Q15 | Já faltei à escola ou deixei de fazer lição de casa por ter que ajudar em atividades de casa (cozinhar, limpar, cuidar de irmãos etc). |
| | | Q31 | Já aconteceu de eu não poder continuar na escola por causa de uma gravidez, seja minha ou da minha parceira. |
| | | Q21 | Já pensei em abandonar a escola por ter tido e/ou ter um problema de saúde na família. |

Imagem 2. Destaque dimensão E-FAM

4. Desenvolvimento

4.3 Análise Sociodemográfica

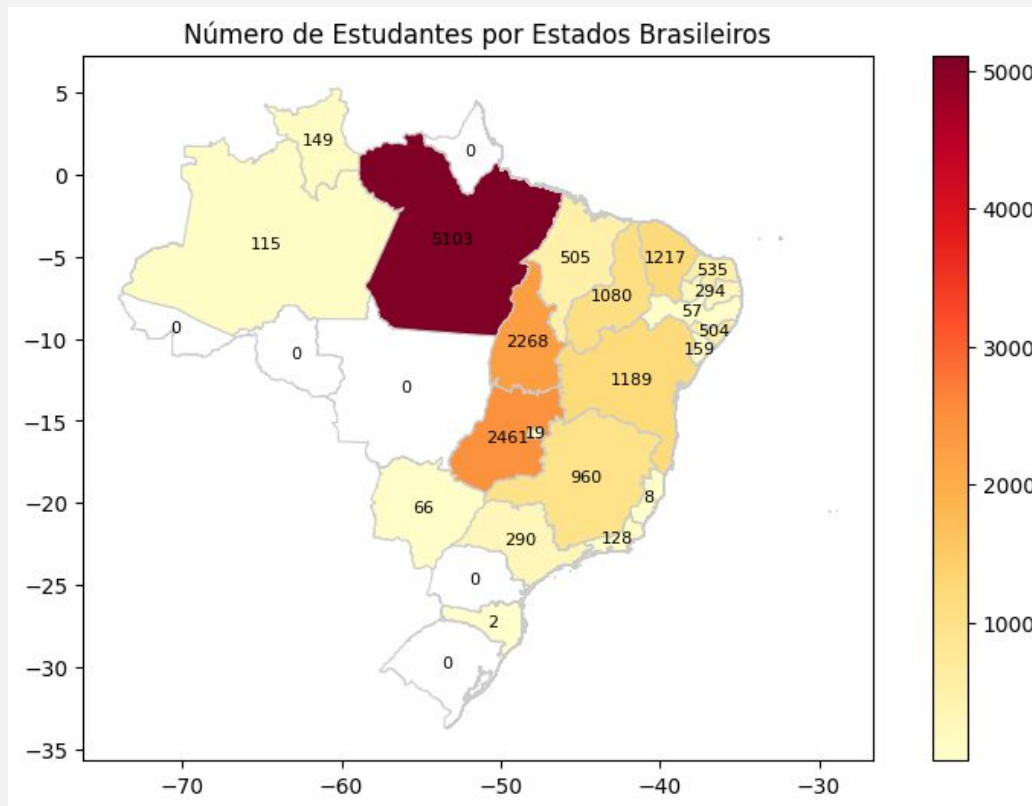


Imagem 3. Distribuição de estudantes por estado

4. Desenvolvimento

4.3 Análise Sociodemográfica

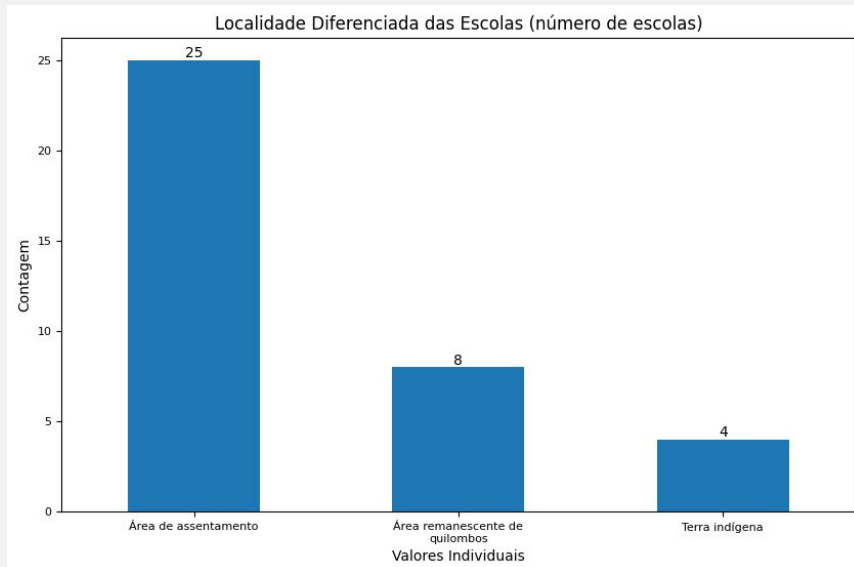


Imagem 4. Número de escolas com localização diferenciada do IAFREE

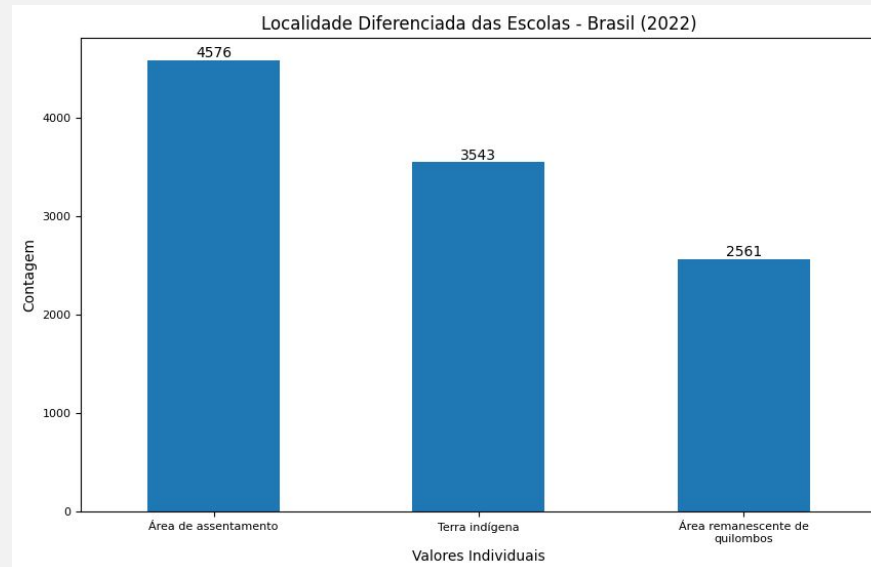


Imagem 5. Número de escolas com localização diferenciada do Censo 2022

4. Desenvolvimento

4.3 Análise Sociodemográfica

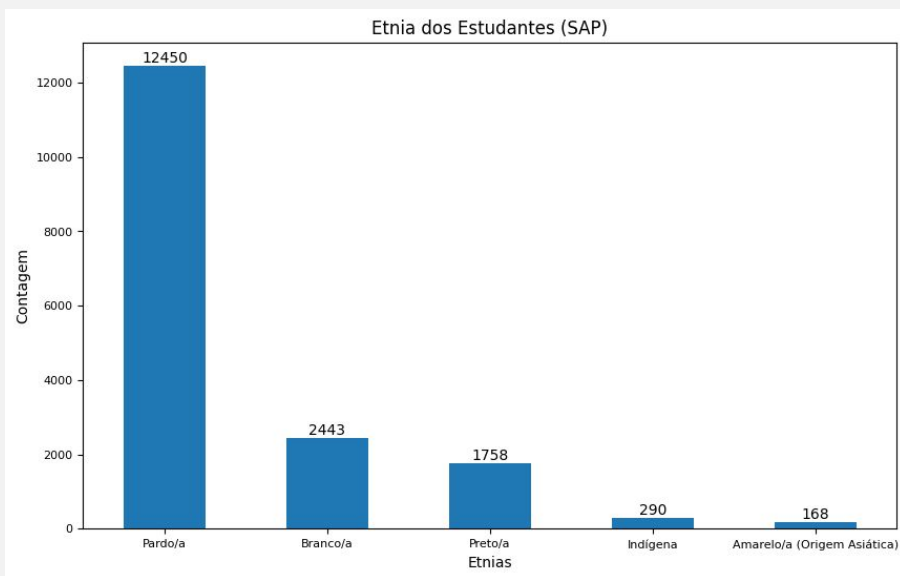


Imagem 6. Etnia dos estudantes que responderam o IAFREE

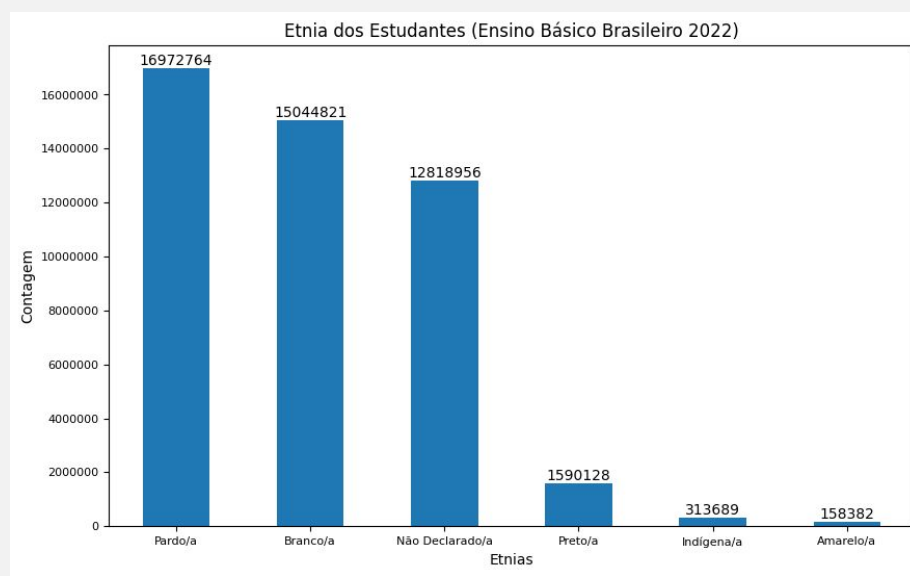


Imagem 7. Etnia dos estudantes segundo o Censo 2022

4. Desenvolvimento

4.4 Análises Estatísticas

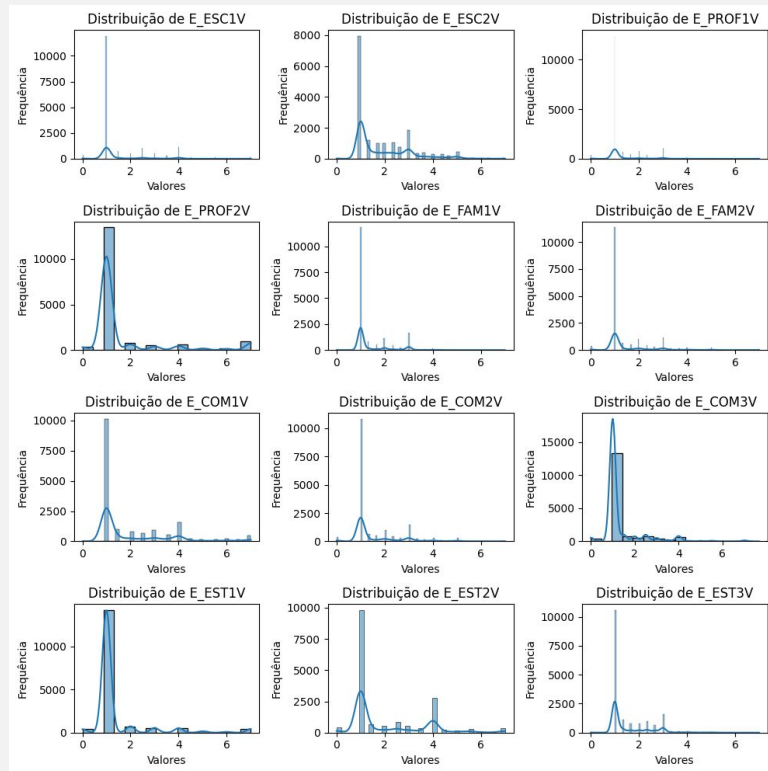
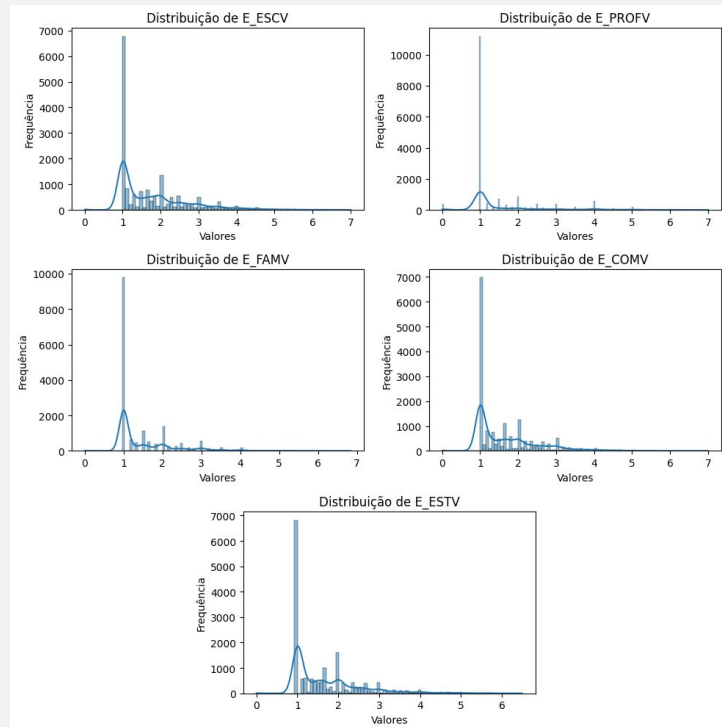



Imagem 8. Histograma de distribuição de dimensões (esquerda) e fatores (direita)

4. Desenvolvimento

4.4 Análises Estatísticas



| Dimensão | Média | Desvio Padrão | Mediana | Mínimo | Máximo |
|----------|----------|---------------|----------|--------|----------|
| E_ESCV | 1.764666 | 0.958435 | 1.416667 | 0.0 | 7.000000 |
| E_PROFV | 1.504817 | 1.067959 | 1.000000 | 0.0 | 7.000000 |
| E_FAMV | 1.489781 | 0.784556 | 1.000000 | 0.0 | 6.833333 |
| E_COMV | 1.677947 | 0.873904 | 1.333333 | 0.0 | 7.000000 |
| E_ESTV | 1.671151 | 0.862080 | 1.333333 | 0.0 | 6.500000 |


Tabela 2. Análises estatísticas dimensões

| Fator | Média | Desvio Padrão | Mediana | Mínimo | Máximo |
|----------|----------|---------------|----------|--------|--------|
| E_ESC1V | 1.567469 | 1.187482 | 1.000000 | 0.0 | 7.0 |
| E_ESC2V | 1.926033 | 1.194067 | 1.333333 | 0.0 | 7.0 |
| E_PROF1V | 1.387515 | 0.928115 | 1.000000 | 0.0 | 7.0 |
| E_PROF2V | 1.622119 | 1.603368 | 1.000000 | 0.0 | 7.0 |
| E_FAM1V | 1.431537 | 0.788930 | 1.000000 | 0.0 | 7.0 |
| E_FAM2V | 1.520058 | 1.062425 | 1.000000 | 0.0 | 7.0 |
| E_COM1V | 2.011573 | 1.588832 | 1.000000 | 0.0 | 7.0 |
| E_COM2V | 1.596665 | 1.117765 | 1.000000 | 0.0 | 7.0 |
| E_COM3V | 1.349514 | 0.963546 | 1.000000 | 0.0 | 7.0 |
| E_EST1V | 1.390223 | 1.224664 | 1.000000 | 0.0 | 7.0 |
| E_EST2V | 2.033257 | 1.571482 | 1.000000 | 0.0 | 7.0 |
| E_EST3V | 1.525026 | 0.844181 | 1.000000 | 0.0 | 7.0 |

Tabela 3. Análises estatísticas fatores

4. Desenvolvimento

4.4 Análises Estatísticas



| | sexo | uf | localizacao | ano_turma | localidade_diferenciada | dependencia_administrativa | id_raca_etnia | porte_da_escola | outras_ofertas_educacionais | id_renda_familiar |
|---------|--------------|--------------|--------------|--------------|-------------------------|----------------------------|---------------|-----------------|-----------------------------|-------------------|
| E_ESCV | 0.039 | 0.938 | 0.486 | 0.372 | 0.066 | 0.568 | 0.141 | 0.000 | 0.201 | 0.000 |
| E_PROFV | 0.000 | 0.141 | 0.010 | 0.025 | 0.579 | 0.000 | 0.951 | 0.041 | 0.000 | 0.150 |
| E_FAMV | 0.000 | 0.075 | 0.000 | 0.000 | 0.000 | 0.050 | 0.028 | 0.000 | 0.000 | 0.000 |
| E_COMV | 0.000 | 0.000 | 0.000 | 0.000 | 0.002 | 0.000 | 0.002 | 0.000 | 0.239 | 0.000 |
| E_ESTV | 0.213 | 0.030 | 0.538 | 0.668 | 0.850 | 0.514 | 0.844 | 0.249 | 0.411 | 0.594 |

Tabela 2. P Valor estatísticas dimensões

| | sexo | uf | localizacao | ano_turma | localidade_diferenciada | dependencia_administrativa | id_raca_etnia | porte_da_escola | outras_ofertas_educacionais | id_renda_familiar |
|----------|--------------|--------------|--------------|--------------|-------------------------|----------------------------|---------------|-----------------|-----------------------------|-------------------|
| E_ESC1V | 0.011 | 0.079 | 0.465 | 0.305 | 0.416 | 0.000 | 0.020 | 0.748 | 0.083 | 0.032 |
| E_ESC2V | 0.001 | 0.064 | 0.109 | 0.199 | 0.138 | 0.009 | 0.935 | 0.000 | 0.002 | 0.000 |
| E_PROF1V | 0.533 | 0.177 | 0.015 | 0.061 | 0.116 | 0.000 | 0.385 | 0.713 | 0.443 | 0.000 |
| E_PROF2V | 0.000 | 0.116 | 0.383 | 0.236 | 0.717 | 0.000 | 0.369 | 0.706 | 0.000 | 0.079 |
| E_FAM1V | 0.000 | 0.159 | 0.093 | 0.000 | 0.000 | 0.000 | 0.004 | 0.000 | 0.000 | 0.065 |
| E_FAM2V | 0.001 | 0.612 | 0.000 | 0.803 | 0.024 | 0.000 | 0.627 | 0.865 | 0.245 | 0.000 |
| E_COM1V | 0.000 | 0.000 | 0.000 | 0.330 | 0.000 | 0.000 | 0.641 | 0.415 | 0.000 | 0.000 |
| E_COM2V | 0.000 | 0.074 | 0.701 | 0.000 | 0.500 | 0.182 | 0.285 | 0.276 | 0.820 | 0.835 |
| E_COM3V | 0.095 | 0.781 | 0.000 | 0.240 | 0.938 | 0.000 | 0.339 | 0.527 | 0.025 | 0.648 |
| E_EST1V | 0.000 | 0.988 | 0.000 | 0.215 | 0.307 | 0.000 | 0.804 | 0.028 | 0.632 | 0.002 |
| E_EST2V | 0.000 | 0.000 | 0.060 | 0.000 | 0.332 | 0.001 | 0.127 | 0.409 | 0.003 | 0.000 |
| E_EST3V | 0.069 | 0.000 | 0.001 | 0.000 | 0.009 | 0.001 | 0.033 | 0.015 | 0.006 | 0.057 |

Tabela 3. P Valor estatísticas fatores

4. Desenvolvimento

4.4 Análises Estatísticas

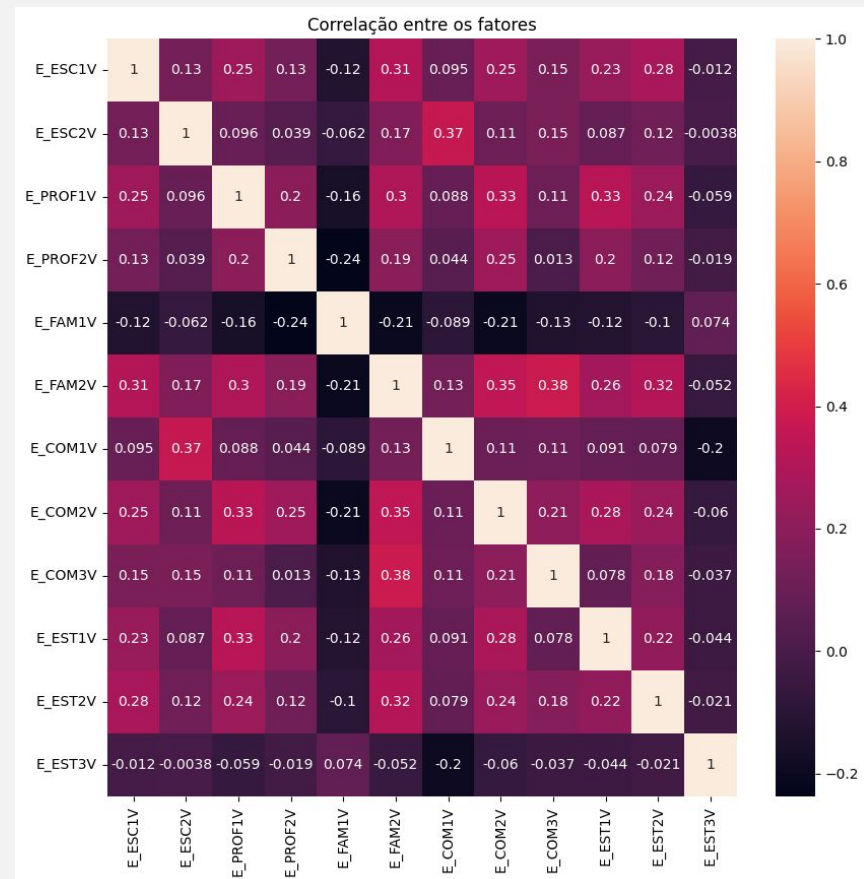
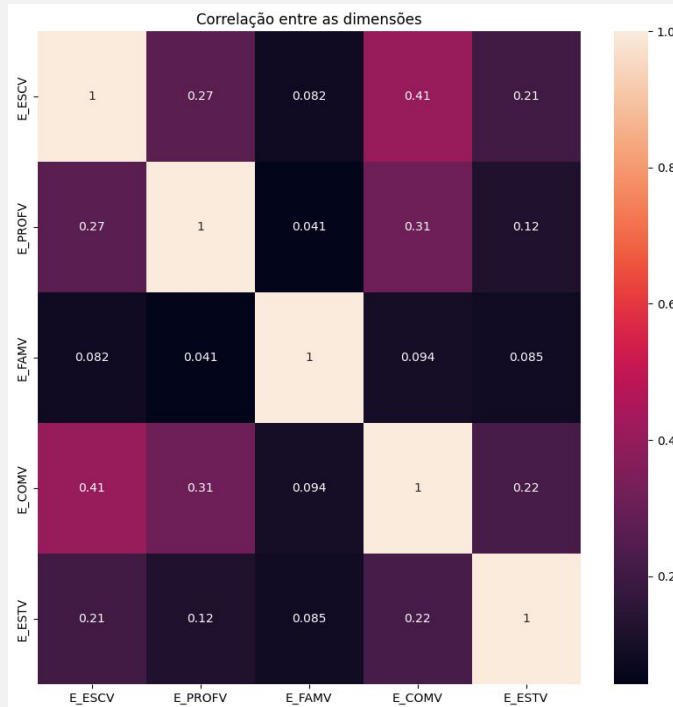


Imagem 9. Correlações entre dimensões (esquerda) e fatores (direita)

4. Desenvolvimento

4.5 Análises Classificadoras



- Desbalanceamento da base tratado ao duplicar linhas;
- Valores reais tratados como inteiros, arredondado para cima;
- 20% da base aleatoriamente escolhida para teste;
- Utilizado **Python 3.10**, biblioteca `sklearn`. Modelos: `DecisionTreeClassifier`, `RandomForestClassifier`, `AdaBoostClassifier`, `CategoricalNB` e `KNeighborsClassifier`;
- Visualização das **matrizes de confusão** para melhor comparação.

4. Desenvolvimento

4.5 Análises Classificadoras



- Melhor performance de algoritmos: **Decision Tree** e **Random Forest**.
 - Acurácia: 74% a 84% em dimensões, 95% a 99% em fatores
 - Precisão: 73% a 84% em dimensões, 95% a 99% em fatores
 - Recall: 74% a 84% em dimensões, 95% a 99% em fatores
 - F1: 73% a 84% em dimensões, 95% a 99% em fatores
- AdaBoost teve o pior desempenho, com 27% a 52% em dimensões e fatores.
- Melhores variáveis para classificação: E_FAMV, para dimensão, e E_FAM1V e E_COM3V, para fatores.



5. Resultados

- Foi investigado a evasão escolar no contexto do Sistema Alerta Preventivo (SAP), empregando técnicas de classificação de dados para identificar fatores de risco mais proeminentes;
- Realizados estudos sociodemográficos, estatísticos e classificatórios para estudo aprofundado da base;
- Trabalhos futuros: análise das perguntas individuais, estudar possíveis motivações associadas às perguntas, estudar medidas de retenção;
- Publicação do artigo: “Sistemas de frequência escolar: avanços, desafios e atuais perspectivas” (WAPLA 2023, evento associado ao CBIE).

6. Referências

- AMO, Daniel et al. Local technology to enhance data privacy and security in educational technology. *International Journal of Interactive Multimedia and Artificial Intelligence*, Universidad Internacional de La Rioja, v. 7, n. 2, p. 262, 2021.
- KOÇ, Mustafa. Learning analytics of student participation and achievement in online distance education: A structural equation modeling. *Educational Sciences: Theory & Practice*, Eğitim Danışmanlığı ve Araştırmaları (EDAM), 2017.
- QUEIROGA, Emanuel Marques et al. A learning analytics approach to identify students at risk of dropout: A case study with a technical distance education course. *Applied Sciences*, MDPI, v. 10, n. 11, p. 3998, 2020.
- LEE, Sunbok; CHUNG, Jae Young. The machine learning-based dropout early warning system for improving the performance of dropout prediction. *Applied Sciences*, MDPI AG, v. 9, n. 15, p. 3093, jul. 2019. ISSN 2076-3417.
- NIYOGISUBIZO, Jovial et al. Predicting student's dropout in university classes using two-layer ensemble machine learning approach: A novel stacked generalization. *Computers and Education: Artificial Intelligence*, Elsevier, v. 3, p. 100066, 2022.



Obrigada

**UDESC – Universidade do Estado de
Santa Catarina**

daniellavasconc@gmail.com

Análises classificatórias sobre os fatores de risco de evasão escolar

Daniella Martins Vasconcellos

Isabela Gasparini
Orientadora

Elaine Harada Teixeira de Oliveira
(UFAM)
Coorientadora

30/11/2023
Trabalho de Conclusão de Curso 2

02

Reconhecimento de padrões

$$\text{Acurácia} = \frac{\text{Verdadeiros Positivos} + \text{Verdadeiros Negativos}}{\text{Total de Instâncias}}$$

Fórmula 1. Cálculo da Acurácia

$$\text{Precisão} = \frac{\text{Verdadeiros Positivos}}{\text{Verdadeiros Positivos} + \text{Falsos Positivos}}$$

Fórmula 2. Cálculo da Precisão

$$\text{Recall} = \frac{\text{Verdadeiros Positivos}}{\text{Verdadeiros Positivos} + \text{Falsos Negativos}}$$

Fórmula 3. Cálculo do Recall

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Fórmula 4. Cálculo do F1 Score

4. Desenvolvimento

4.3 Análises Classificadoras

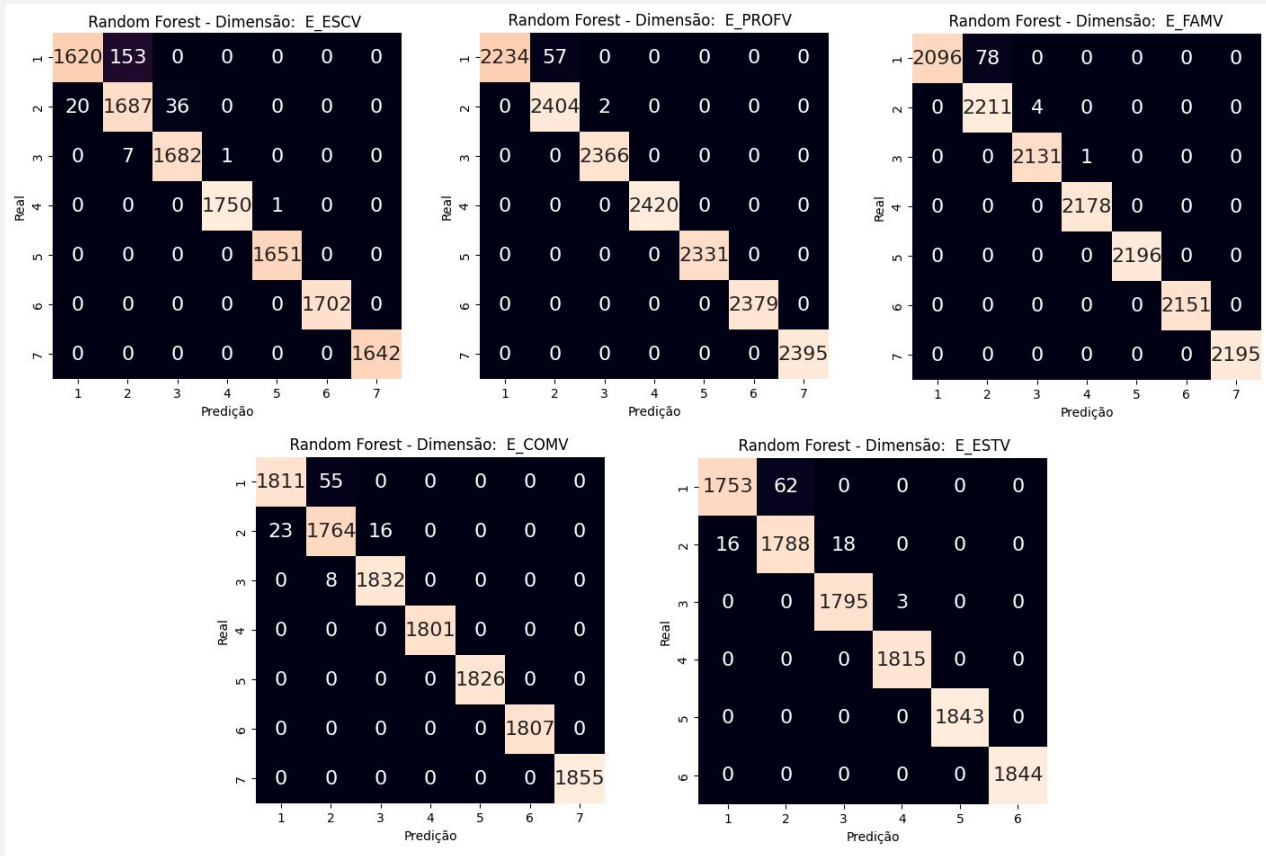


Imagem 10. Matrizes de confusão das dimensões de Random Forest

4. Desenvolvimento

4.3 Análises Classificadoras

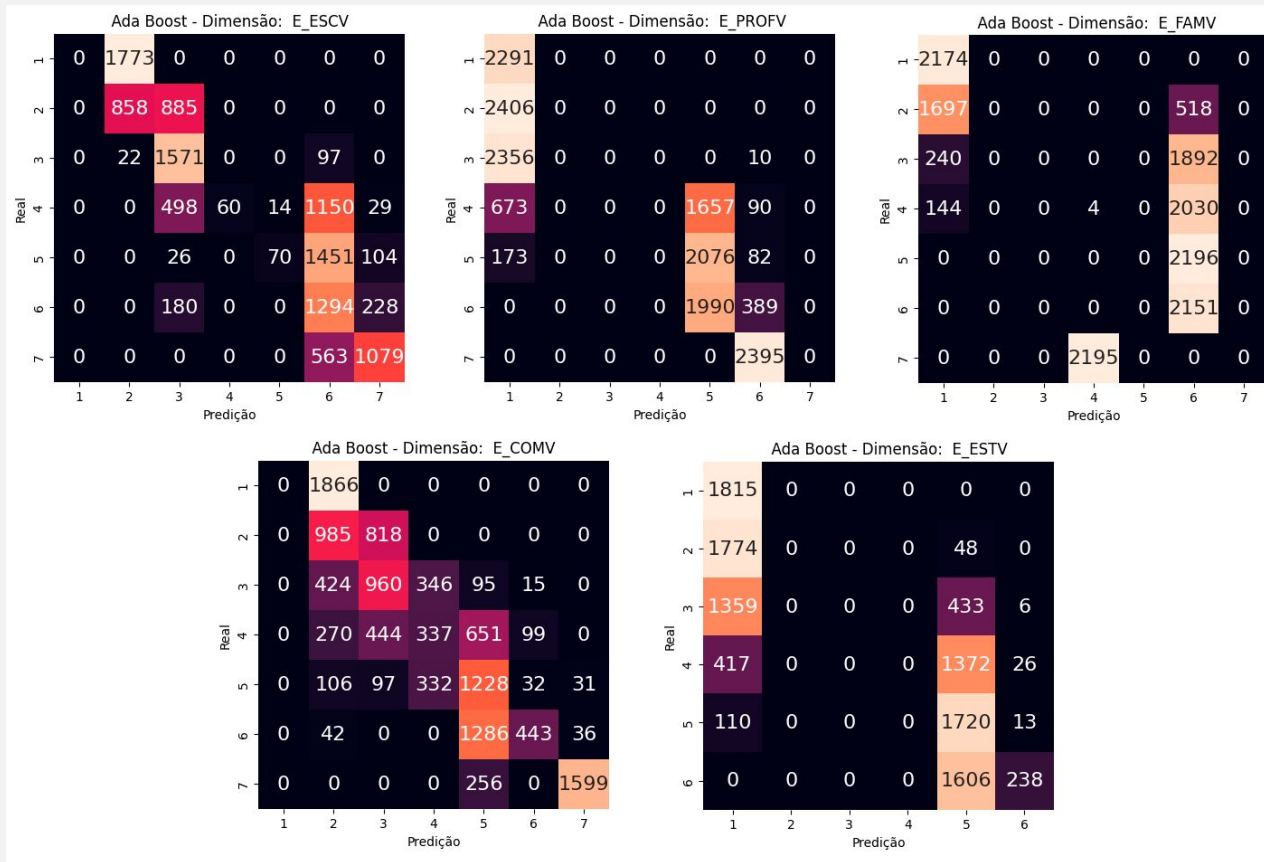


Imagem 11. Matrizes de confusão das dimensões de AdaBoost



3. Trabalhos relacionados

01

A learning analytics approach to identify students at risk of dropout: A case study with a technical distance education course. (QUEIROGA et al, 2020)

02

The machine learning-based dropout early warning system for improving the performance of dropout prediction (LEE e CHUNG, 2019)

03

Predicting student's dropout in university classes using two-layer ensemble machine learning approach: A novel stacked generalization (NIYOGISUBIZO et al, 2022)