

## CLASSIFICAÇÃO DE FATORES QUE PODEM INFLUENCIAR NA PERMANÊNCIA ESCOLAR DOS ALUNOS DOS CURSOS TÉCNICOS INTEGRADOS DO IFRN CAMPUS IPANGUAÇU

**Dimas Lopes de Araújo Junior (Acadêmico do Curso de Licenciatura em Informática, Instituto Federal do Rio Grande do Norte - IFRN)**  
**Anny Kelly Crispim da Silva Lima (Acadêmica do Curso de Licenciatura em Informática, Instituto Federal do Rio Grande do Norte - IFRN)**  
**Fabricio Nascimento de Brito (Acadêmico do Curso de Licenciatura em Informática, Instituto Federal do Rio Grande do Norte - IFRN)**  
**Hugo Wendell Maia (Orientador) Mestre em Ciência da Computação pelo Programa de Pós-Graduação PPgCC da Universidade Federal Rural do Semi-Árido e Universidade do Estado do Rio Grande do Norte – UFRSA**  
**Email: araujo.junior2030@gmail.com, kellyanny067@gmail.com, fabricion974@gmail.com, hugo.maia@ifrn.edu.br**

### 1. INTRODUÇÃO

Nos anos de 2020 e 2021, o mundo sofreu com problemas decorrentes da Covid-19. A evasão escolar é um enfático problema da sociedade. Com a pandemia, a educação sofreu um grande impacto negativo nesse período e ainda está se recuperando dessa adversidade. Para tanto, é preciso elencar e classificar os principais fatores que podem influenciar na evasão, permanência e conclusão dos alunos. Com isso, foi utilizado como amostragem os dados dos alunos dos cursos técnicos integrados de nível médio do IFRN - Campus Ipanguaçu.

### 2. MATERIAIS E MÉTODOS

A metodologia do trabalho foi dividida em três partes, a primeira foi o levantamento da base de dados que desejamos trabalhar. Após isso temos as correções da nossa base de dados, e por último vem a parte do processamento dos dados. Para aplicação da nossa metodologia, utilizamos como base o livro "Inteligência Artificial: uma linguagem de máquina".

Após o recebimento da nossa base de dados, foi realizado o tratamento dos dados inconsistentes ou faltosos, e a seleção das colunas que seriam os nossos previsores e classe. Os previsores escolhidos foram: Ano de Ingresso, Deficiência, Descrição do curso, Etnia/Raça, I.R.A., Renda Per Capta, Tipo de Escola de Origem, Turno e Zona Residencial e a classe foi a Situação no Curso. Foi optado por reduzir as situações de matrícula de 8 para apenas 4, a fim de facilitar nossa classificação. As situações de matrícula "Cancelado", "Cancelamento Compulsório" e "Jubilado" foram classificadas apenas como "Cancelado". As situações "Matrícula Vínculo Institucional", "Matriculado" e "Transferido Interno" foram classificadas apenas como "Matriculado". As situações "Concluído" e "Evasão" permaneceram inalteradas.

Feita a seleção dos previsores e da classe, e corrigidas as inconsistências presentes, o formato de arquivo da base de dados foi convertido de .XLS para .CSV, para facilitar a utilização da mesma no aplicativo Google Colab. O app foi escolhido devido sua praticidade e ser uma ferramenta online e gratuita, que possibilita o manuseio da nossa base de dados utilizando a linguagem de programação Python.

### 3. RESULTADOS E DISCUSSÃO

Primeiro é feita a análise dos resultados obtidos no período pré-pandêmico (2017 a 2019). Nesse período os resultados de ganho de informação de cada variável podem ser observados na tabela "1".

Variável	Ganho de informação
Ano de Ingresso	0.28399402
Deficiência	0.01300609
Descrição do Curso	0.01924633
Etnia/Raça	0.04025875
I.R.A.	0.50519321
Renda Per Capita	0.09412833
Sexo	0.01084807
Tipo de Escola de Origem	0.01001489
Turno	0.01298763
Zona Residencial	0.01032268

Tabela 1: resultados obtidos no período de 2017 a 2019

Quanto maior o ganho de informação, mais influência essa variável tem no resultado final (RUSSEL, 2013), portanto fazendo uma análise dos resultados obtidos na tabela podemos afirmar que a variável mais influente na classificação dos alunos do período foi a variável "I.R.A.", seguida do "Ano de ingresso" e "Renda per capita", enquanto a variável menos influente foi "Tipo de escola de origem", seguida por "Zona residencial" e "Sexo".

Variável	Ganho de informação
Ano de Ingresso	0
Deficiência	0.02696682
Descrição do Curso	0.02696682
Etnia/Raça	0.07915797
I.R.A.	0.51215622
Renda Per Capita	0.14198157
Sexo	0.0859615
Tipo de Escola de Origem	0.01218806
Turno	0.01957744
Zona Residencial	0.09504362

Tabela "2": Resultados obtidos nos anos de 2020 e 2021

As variáveis "I.R.A." e "Renda per capita" aparecem entre os fatores mais influentes na classificação dos alunos, porém dessa vez a variável "Ano de ingresso" está com ganho de informação igual a zero, mesmo após diversos treinamentos realizados no aplicativo Google Colab. As variáveis que obtiveram os menores ganhos de informação foram "Ano de ingresso", "Tipo de escola de origem" e "turno".

Os resultados obtidos indicam que os previsores renda per capita, zona residencial e sexo obtiveram um aumento no ganho de informação durante a pandemia em relação aos anos anteriores analisados.

### 4. CONCLUSÃO

É de profunda relevância a análise cautelosa dos dados que resultam na evasão escolar. Com o tratamento desses dados através do uso da tecnologia de inteligência artificial, é possível ter a tentativa de prever futuros problemas no cenário da educação, levando dados que possam contribuir para métodos preventivos de evasão e contribuir para o êxito na formação do aluno.

### 5. REFERÊNCIAS

FACELI, Katti et al. INTELIGÊNCIA ARTIFICIAL. – UMA ABORDAGEM DE APRENDIZADO DE MÁQUINA. 2ª Edição – Editora LTC – Rio de Janeiro, 2023.

RUSSEL, Stuart; NORVIG, Peter. INTELIGÊNCIA ARTIFICIAL. Elsevier Editora Ltda. tradução Regina Célia Simille. – Rio de Janeiro: Elsevier, 2013.

ROCHA, Tacia. Inteligência Artificial, educação e trabalho: entrevista com Eric Aislan Antonelo. **Texto Livre**, 2019