

**MBA
USP
ESALQ**

O Ensino Médio Paulista e as variáveis desprezadas na influência do desempenho no IDESP entre 2018 e 2022

**Luiz Paulo Rodrigues Almeida
Patrícia Belfiore Fávero**

Introdução:

Panorama da rede pública de ensino básico de São Paulo



AI x AF x EM



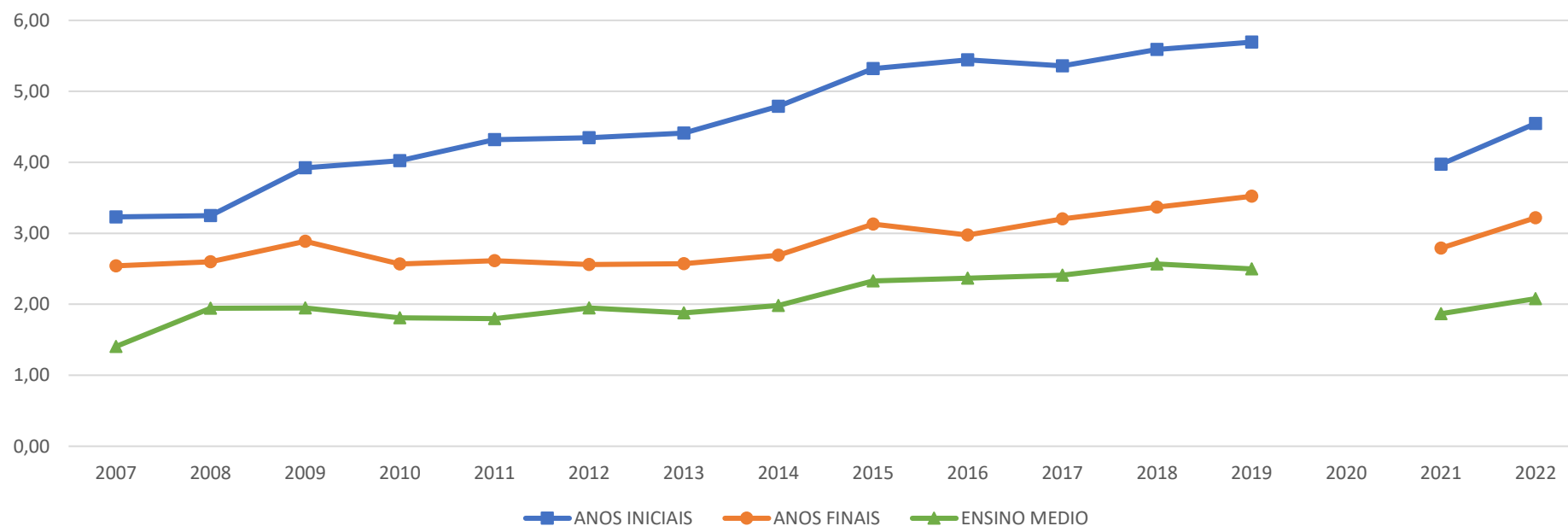
Avaliações



Indicadores

O governo definiu metas de longo prazo, para cada ciclo, a serem atingidas até 2030. Utilizando a pontuação do IDESP para o acompanhamento dessas metas, são elas: **7** para os **Anos Iniciais**, **6** para os **Anos Finais** e **5** para o **Ensino Médio**.

No entanto tais metas estão longe de serem atingidas, principalmente para o Ensino Médio.



Como aperfeiçoar o ensino para que as metas sejam atingidas?

Antes de direcionar recursos públicos, saber onde aloca-los é fundamental.

Nesse contexto, além dos fatores já conhecidos que exercem maior influência no desempenho dos alunos, como a Escolaridade dos Pais e o Índice de Nível Socioeconômico, ligadas mais aos alunos do que aos docentes, talvez existam outros fatores que possam influenciar diretamente o aprendizado.



Variáveis comumente desprezadas

Esse estudo tem por objetivo identificar se diversas variáveis, não ligadas diretamente aos alunos, mas sim aos docentes e às escolas tem o poder de influenciar os resultados de cada escola no IDESP.

Ao lado temos algumas dessas variáveis:

Descrição
Escola recebe o Adicional de Local de Exercício
Obtenção do Bônus por resultado no ano anterior
Média de anos de experiência dos docentes por escola
Percentual de Mestres/Doutores por escola
Percentual de Pós-graduados por escola
Média de alunos por turma por escola
Escola aderente ao Programa Ensino Integral
Percentual de ausência dos professores por escola
Percentual de professores estáveis por escola
Percentual de docentes que obtiveram evolução funcional
Percentual de professores regulares, intermitentes, ingressantes ou que abandonaram a carreira

Materiais e métodos:

Essa pesquisa usou **dados secundários**, dados disponibilizados através do portal Dados Abertos da Educação.



Acompanhamento
de Planos e
Programas



Infraestrutura e
Serviços



Matrículas na
Rede



Orçamento da
Educação



Profissionais da
Educação



Resultados
Educacionais

Criada em 2018, a plataforma contém dados de mais de **5 mil escolas**, cerca de **250 mil profissionais da educação** e dos mais de **1,4 milhão de alunos** da rede pública estadual além do resultados dos exames aplicados aos alunos para avaliação educacional.

Análise exploratória e ajustes



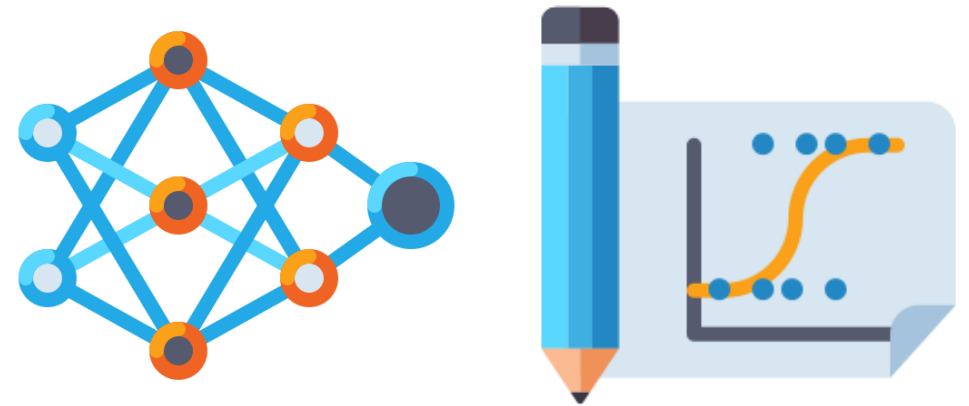
Explorando e ajustando os dados com **RStudio** e **BigQuery**

- Grande volume de dados
- Arquivos mensais e anuais
- Inúmeras inconsistências
- Filtrar apenas o QM e EM, **150 mil docentes e 3,5 mil escolas**
- Consolidação dos dados
- Insights interessantes

Aplicação de modelos

Foram utilizados modelos de **Regressão Logística** e de **Redes Neurais Artificiais** a fim de verificar quais variáveis possuem significância estatística.

Para avaliar o desempenho dos modelos foi utilizada uma **matriz de confusão** onde se extraíram métricas derivadas da matriz, além da AUC da curva ROC e do “LogLikelihood”.

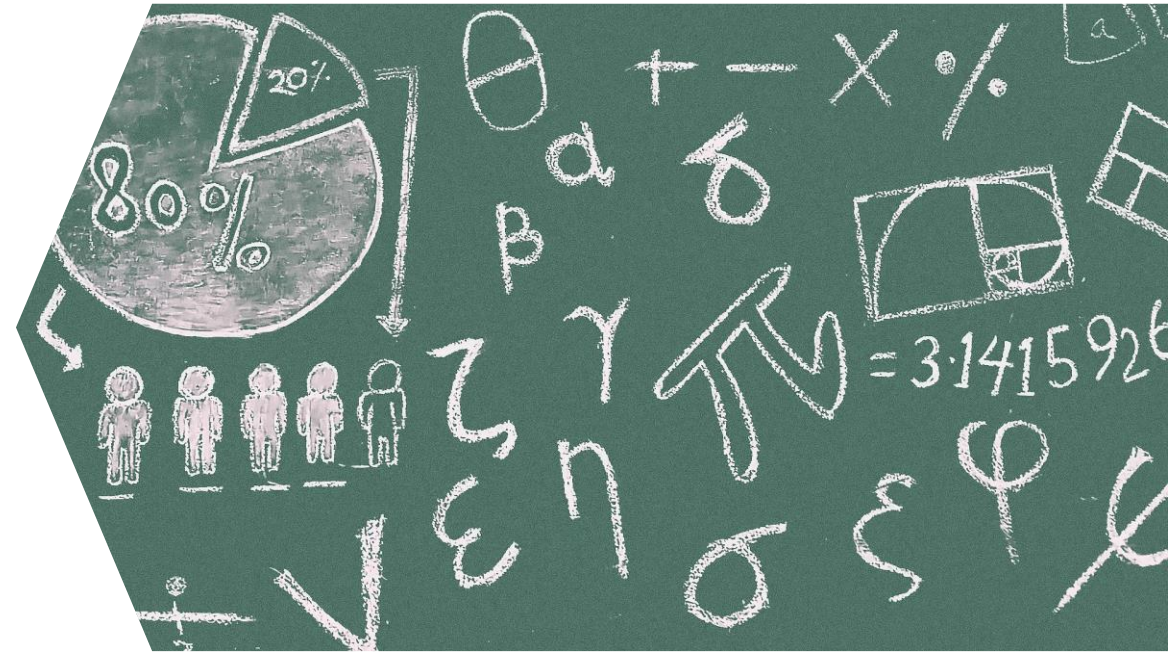


		Valores reais	
		Negativos	Positivos
Valores Preditos	Negativos	Verdadeiros Negativos (VN)	Falsos Negativos (FN)
	Positivos	Falsos Positivos (FP)	Verdadeiros Positivos (VP)

Resultados e discussão:

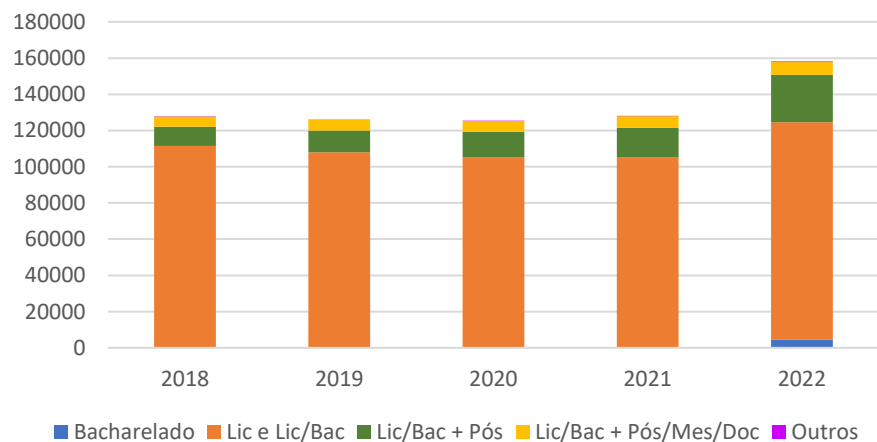
Além dos modelos em si, diversos insights podem ser extraídos das informações analisadas.

- Rotatividade dos professores
- Formação dos docentes
- Ausências dos profissionais
- Crescimento das escolas PEI
- Impacto de políticas públicas
- Mudança das categorias

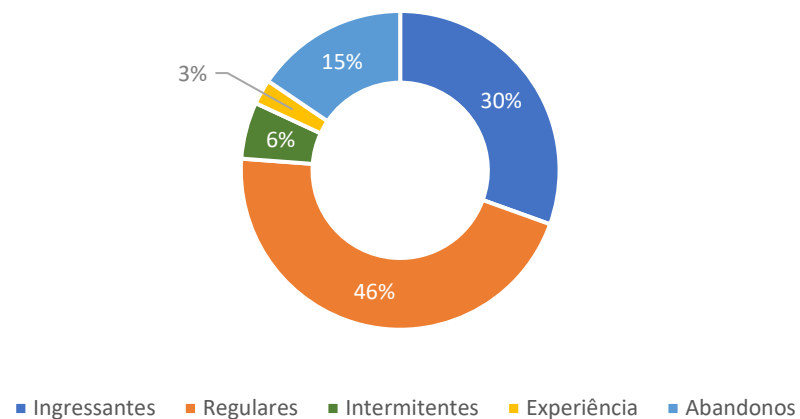


Insights sobre os docentes

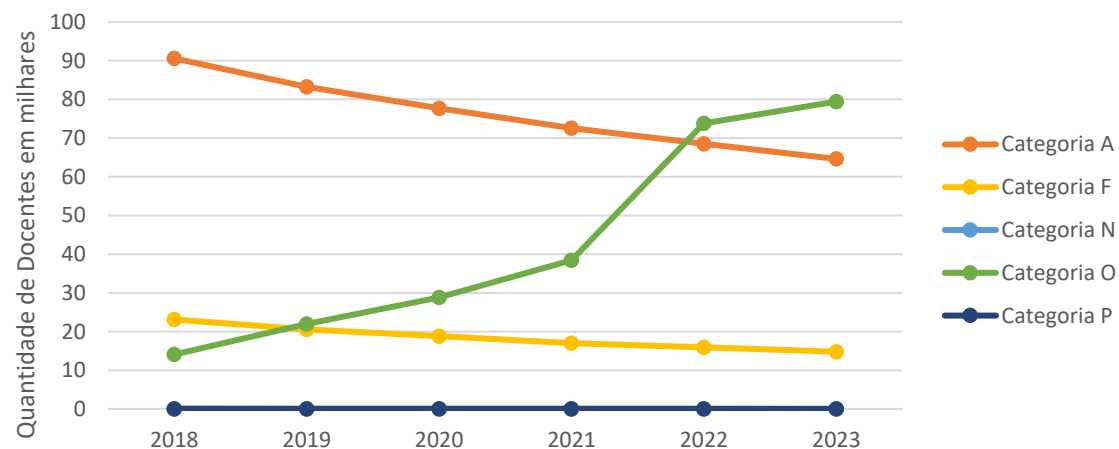
Formação dos Professores - EM - 2018 a 2022



Rotatividade de Professores - EM - 2018 a 2022



Categoria dos Docentes – EM – 2018 a 2023



Insights sobre os docentes e as escolas

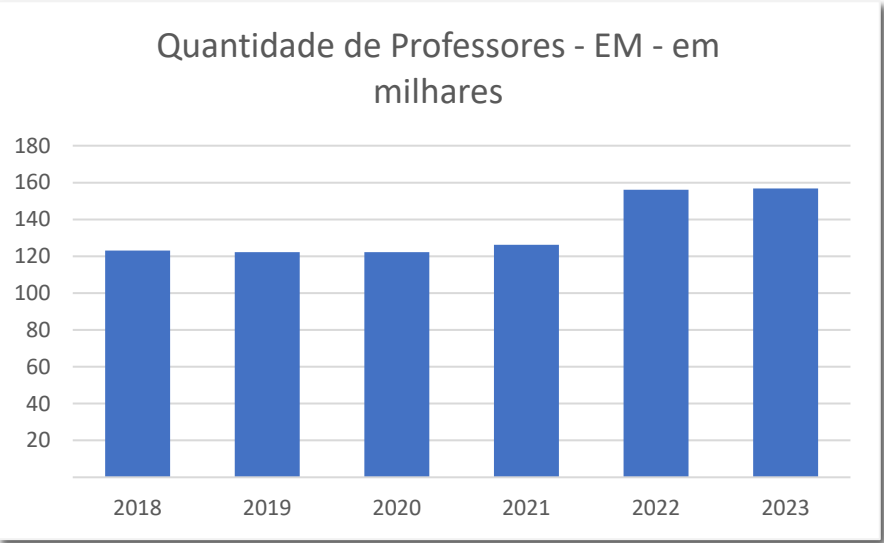


Tabela 1. Ausências de Professores do EM da rede pública de SP

Dados	2018*	2019	2020**	2021	2022	2023
Mín. Faltas/Esc	4	2	1	3	7	5
Máx. Faltas/Esc	1630	6091	4371	6126	6912	8662
Média Faltas/Esc	400	1386	708	956	1196	1270
Média Faltas/Prof	11,53	40,20	20,54	26,85	27,16	28,73
Total Faltas	1419017	4915497	2511340	3389258	4241334	4503644

Fonte: Dados Abertos da Educação

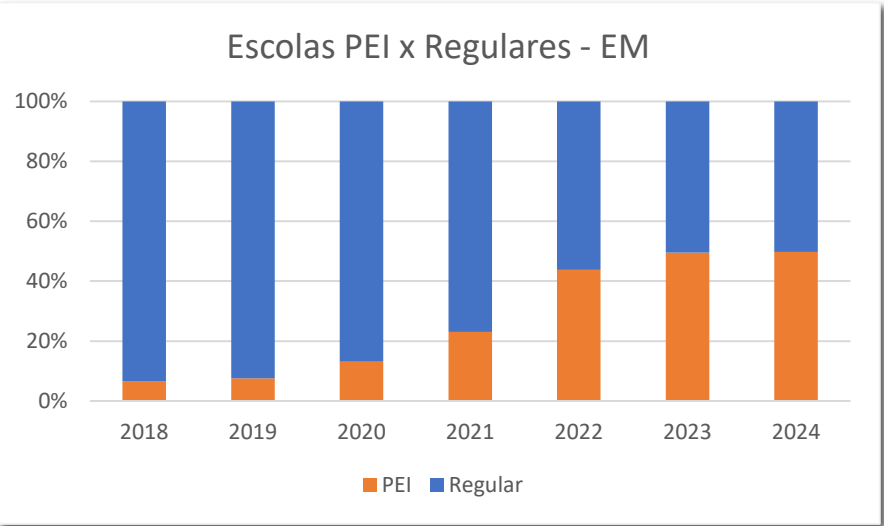


Tabela 2. Escolas do EM que atingiram a meta do IDESP –PEI x Regulares

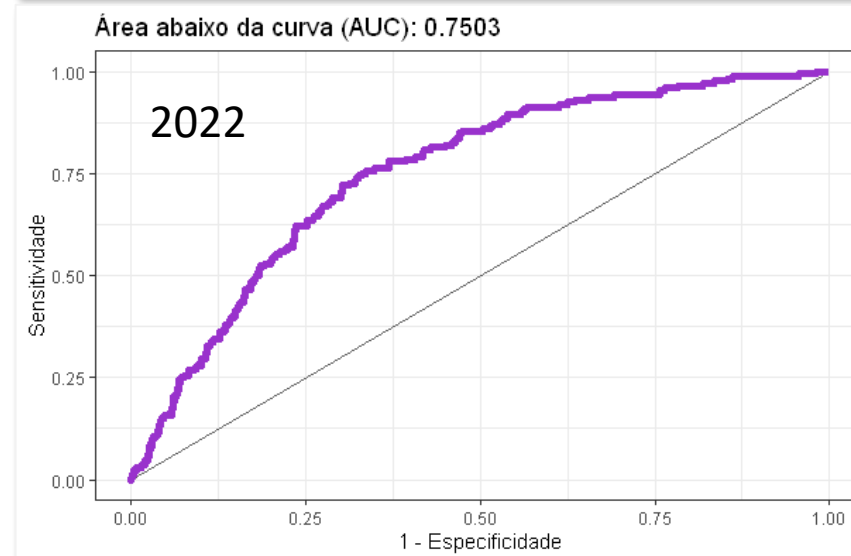
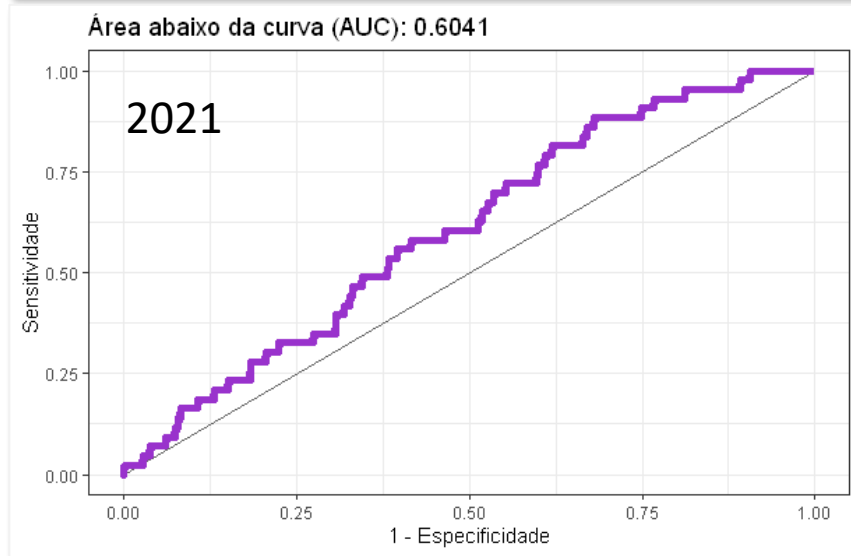
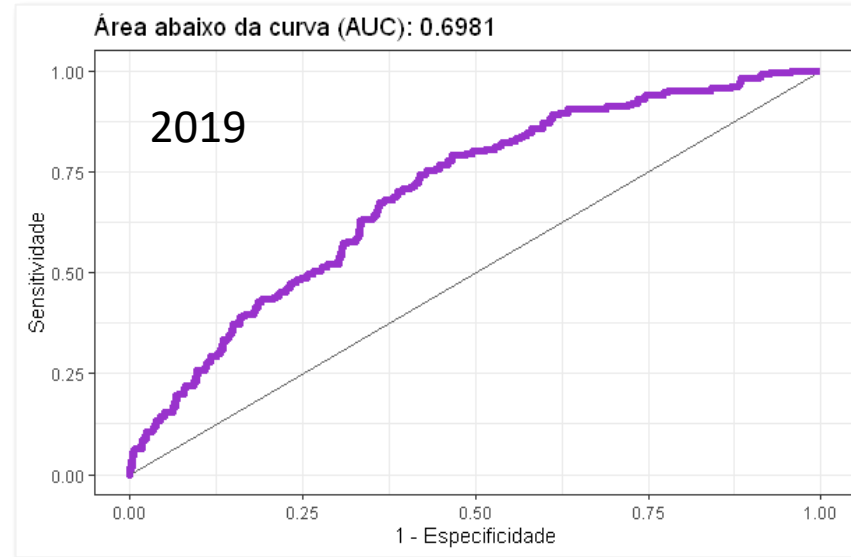
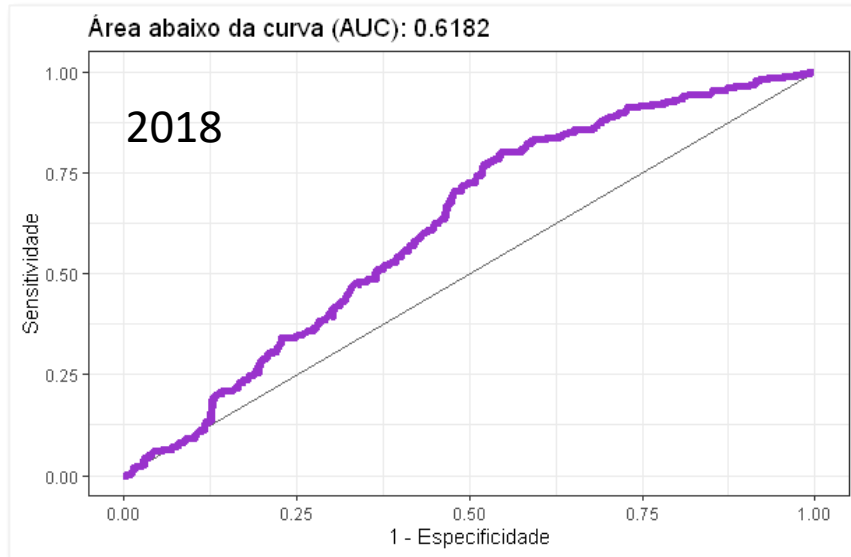
Modelo de Ensino	2018	2019	2021	2022
PEI	57,80%	46,77%	10,05%	35,46%
Regular	41,95%	25,87%	6,05%	11,13%

Fonte: Dados Abertos da Educação

* Dados disponíveis apenas dos meses de Outubro, Novembro e Dezembro

** Dados do mês de Maio não disponíveis

Curvas ROC do modelo de Regressão Logística



Curvas ROC para os modelos de Regressão Logística

Variáveis preditoras, sua significância, magnitude e sinal

Tabela 3. Variáveis preditoras e seus valores estimados após procedimento Stewise

Variáveis	2018	2019	2021	2022	18 a 22
Intercepto	0,32817	0,977976	-2,144778	-0,125462	-0,723381
ALE	-	-	-	-	-0,143629
BONUS	-1,05325	-0,936754	-	-1,573165	-1,337081
ANOS EXPERIENCIA	-0,03378	-	-0,194735	-	-0,121550
% MESTRES/DOUTORES	-	-	0,039831	-0,015146	-
% POS GRADUADOS	-	0,008316	-0,045923	-	-0,003460
MEDIA ALUNOS/TURMA	-	-0,021216	-	-0,067577	-0,039896
PEI	0,92303	0,849212	0,644951	0,893310	0,267427
% AUSENCIA DOCENTE	-	-0,039742	-	-0,026877	-0,056947
% PROF ESTAVEIS	-	-0,009815	0,009248	0,005650	0,036746

Fonte: Resultados originais da Pesquisa

< -Negativamente significativo

Positivamente significativo ->

Comparativo entre os modelos

Tabela 4. Métricas do modelo de Regressão Logística

Modelo	Acurácia	Recall	Especificidade	F1 Score	AUC
Regressão logística - 2018	57,57%	48%	65%	49%	62%
Regressão logística - 2019	72,52%	8%	98%	14%	70%
Regressão logística - 2021	78,14%	23%	82%	15%	60%
Regressão logística - 2022	75,74%	10%	97%	16%	75%
Regressão logística - 18 a 22	76,42%	18%	95%	25%	74%

Fonte: Resultados originais da Pesquisa

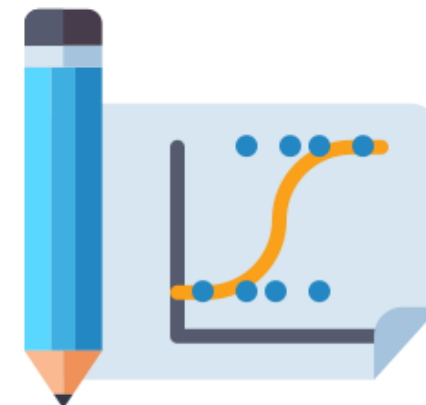
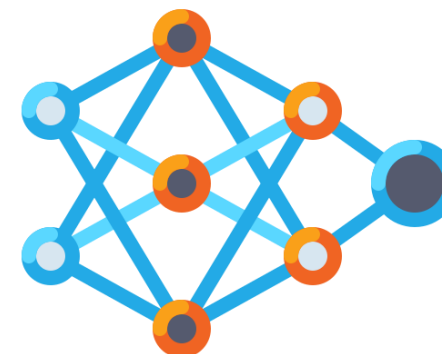


Tabela 5. Métricas do modelo de Redes Neurais Artificiais

Modelo	Acurácia	Recall	Especificidade	F1 Score	AUC
Rede neural - 2018	56,43%	49%	62%	51%	62%
Rede neural - 2019	72,24%	61%	73%	13%	71%
Rede neural - 2021	82,09%	11%	95%	16%	62%
Rede neural - 2022	74,75%	38%	76%	11%	75%
Rede neural - 18 a 22	74,75%	37%	76%	11%	74%

Fonte: Resultados originais da Pesquisa



Conclusões:

O desempenho de ambos os modelos, utilizando as variáveis centradas apenas nas questões docentes e estruturais da rede estadual, se mostrou consideravelmente baixo.

- Adição de novas variáveis explicativas, focadas nos alunos, podem criar um modelo mais robusto;
- O sinal e magnitude das variáveis indicam quais delas tem maior influência;
- A influência das políticas públicas é visível ao analisar os dados e cruzá-los com os modelos.

Agradecimentos:

Agradeço à minha parceira Elisa por acreditar em mim e à minha irmã Ana Cláudia que sempre me arranca uma risada quando tudo está tenso e sério.

Agradeço também à professora Patrícia Belfiore pela orientação nesse trabalho.



"Se cheguei até aqui, foi porque me apoiei no ombro dos gigantes"