

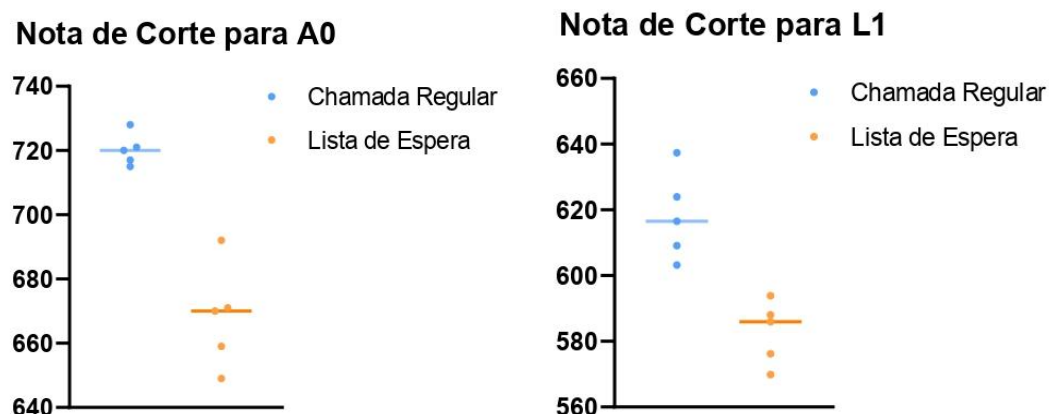


ALÉM DA NOTA DE CORTE: UMA SOLUÇÃO INTERATIVA PARA ACOMPANHAR CONVOCAÇÕES NO SISU

Pedro Santos Rosa Muradás e Vitoria Pires da Silva

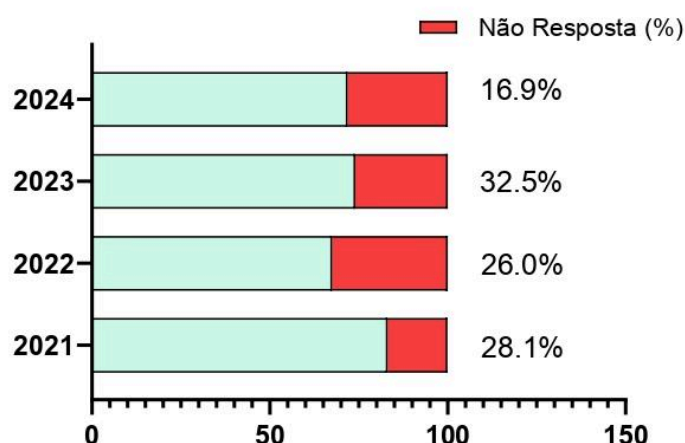
1 Contextualização da análise

A nota de corte divulgada pelo Sistema de Seleção Unificada (SiSU) costuma refletir apenas a nota mínima necessária para ser aprovado na chamada regular. No entanto, essa informação não oferece uma visão completa do processo seletivo, especialmente em universidades como a Universidade Federal do ABC, que realizam diversas chamadas subsequentes por meio da lista de espera. A nota final do último candidato convocado, que pode ser significativamente inferior à da chamada regular, raramente é divulgada de forma acessível e organizada. E assim, muitos candidatos, por não compreenderem essa dinâmica, acreditam que suas chances terminam com a chamada regular e, por consequência, deixam de acompanhar o processo. Isso leva ao abandono precoce do processo seletivo, não necessariamente por falta de interesse, mas por ausência de informação clara e consolidada.



Segundo dados preliminares, estima-se que mais de 3.600 candidatos que passaram na lista de espera deixaram de solicitar matrícula na UFABC nos últimos 4 anos somente, possivelmente por não saberem que ainda tinham chances reais de convocação ou não se atentarem aos chamados via e-mail. A ausência de um panorama completo do processo de seleção e fácil acessibilidade dessa informação leva a uma subutilização das oportunidades disponíveis e impacta diretamente na ocupação de vagas públicas.

Solicitação de Matrícula de pessoas na lista de Espera



Diante deste cenário, propomos a criação de uma ferramenta interativa, que consolide as informações de todas as chamadas da UFABC, indicando, por curso, modalidade de ingresso e nota do ENEM, sua possível colocação e porcentagem de aceitação na inscrição para a universidade dentro da chamada regular ou listas de espera. A ferramenta tem como objetivo aumentar a transparência, orientar melhor os candidatos e evitar que oportunidades sejam desperdiçadas por falta de informação.

2 Dados utilizados

Listas de convocação da UFABC por curso e chamada que foram disponibilizados.

[bd_prograd01_2024](#)

[bd_prograd01_2023](#)

[bd_prograd01_2022](#)

[bd_prograd01_2021](#)

[bd_prograd01_2020](#)

3 Teoria ou ferramenta utilizada

Para a realização desta análise e construção da ferramenta interativa, foram utilizadas três ferramentas principais: Google Sheets, Google Colab (Python) e o aplicativo GraphPad Prism. Cada uma delas foi escolhida de forma estratégica,

considerando critérios como acessibilidade, eficiência no processamento de dados e qualidade na apresentação visual dos resultados.

O Google Sheets foi adotado para a visualização tanto a nota de corte da chamada regular quanto a nota do último convocado em cada modalidade.

Para o processamento e análise dos dados, foi utilizado o Google Colab, ambiente de desenvolvimento baseado em Python. Essa plataforma permitiu a manipulação eficiente de grandes volumes de dados extraídos das planilhas disponibilizadas pela Pró-Reitoria de Graduação (ProGrad) da UFABC entre os anos de 2020 e 2024. Utilizamos bibliotecas como Pandas, NumPy e Matplotlib para realizar as seguintes tarefas:

- Integração de arquivos com diferentes formatos e estruturas referentes a múltiplos anos e chamadas do SiSU, por meio de operações otimizadas de merge e concatenação de dados;
- Limpeza e padronização das bases, com o tratamento de valores ausentes, correção de inconsistências e uniformização da nomenclatura entre os documentos originais;
- Extração automatizada da nota do último convocado por curso, turno e modalidade, considerando todas as chamadas da lista de espera;
- Aplicação de técnicas básicas de machine learning, como modelos de classificação probabilística, para estimar as chances de convocação com base no histórico de dados, levando em conta nota, curso, modalidade e ano de inscrição.

4 Metodologia

4.1 Metodologia Teórica

A construção da ferramenta interativa de acompanhamento das convocações do SiSU na UFABC foi desenvolvida a partir de uma série de etapas organizadas e integradas, combinando ferramentas acessíveis e métodos computacionais eficientes. Inicialmente, foram coletadas as listas públicas de convocação da UFABC entre os anos de 2020 e 2024, incluindo tanto a chamada regular quanto todas as chamadas da lista de espera. Esses arquivos, organizados sob a nomenclatura padrão `bd_prograd01_ano`, continham as informações necessárias para identificar o comportamento histórico de convocações por curso, turno e modalidade de ingresso.

Com os dados em mãos, utilizamos a linguagem Python por meio do ambiente virtual Google Colab para consolidar, limpar e padronizar as informações. As bibliotecas Pandas, NumPy e sklearn foram fundamentais nesse processo. Realizamos a remoção de duplicatas, correção de inconsistências de nomenclatura e unificação dos diferentes arquivos em uma base de dados única e estruturada, facilitando as análises posteriores. Em seguida, identificamos, por meio de agrupamentos e ordenações, o último candidato convocado em cada curso, modalidade e turno, considerando sempre a última chamada da lista de espera em cada ano.

Utilizamos o Prism para gerar gráficos comparativos entre a nota da chamada regular e a nota do último convocado, bem como distribuições por modalidade e curso, revelando de forma intuitiva os padrões de chamada e as diferenças expressivas de notas.

4.2 Metodologia do código

Neste projeto, desenvolvemos uma ferramenta preditiva baseada em aprendizado de máquina com o objetivo de estimar a probabilidade de um candidato ser convocado em diferentes chamadas do processo seletivo da UFABC, considerando suas informações de nota, cota e curso de interesse. Para isso, utilizamos um conjunto de dados contendo registros históricos de candidatos, incluindo suas notas, modalidade de inscrição e curso escolhido.

Inicialmente, os dados foram pré-processados. As variáveis: modalidade de inscrição e opção de curso — foram codificadas numericamente utilizando a técnica de *Label Encoding*, a fim de torná-las compatíveis com o modelo de aprendizado. A variável alvo (chamada no código de 'y') representa a chamada em que o candidato foi convocado, com a classe 0 para chamada regular, 1, 2, 3 e 4 para listas de espera e 5 indicando não convocado.

Em seguida, foi treinado um modelo de classificação do tipo *Random Forest* (*RandomForestClassifier*), por ser um algoritmo robusto, com bom desempenho em dados tabulares e capaz de lidar com variáveis categóricas e interações entre atributos. Basicamente o *Random Forest* é um modelo de machine learning supervisionado usado tanto para classificação (como prever em qual chamada um

aluno será convocado) quanto para regressão (prever um valor numérico). O modelo foi treinado com 80% dos dados e testado com os 20% restantes, garantindo reprodutibilidade com a definição de uma semente aleatória fixa. Para mitigar o efeito de desbalanceamento entre classes, foi utilizada a opção *class_weight='balanced'*. Por fim, desenvolvemos uma interface interativa permitindo que o usuário insira suas informações e recebe como retorno a probabilidade de ser convocado em cada chamada, além de uma previsão da chamada mais provável.

5 Resultados obtidos

Nota:	690
Modalidade:	A0
Curso:	BCT
Prever	
Probabilidades por chamada:	
Chamada 0: 0.00%	
Chamada 1: 13.78%	
Chamada 2: 60.51%	
Chamada 3: 25.67%	
Chamada 4: 0.00%	
Chamada 5: 0.04%	
Resultado final: Chamada 2	

Nota:	636
Modalidade:	L1
Curso:	BCT
Prever	
Probabilidades por chamada:	
Chamada 0: 0.00%	
Chamada 1: 91.00%	
Chamada 2: 8.00%	
Chamada 3: 0.00%	
Chamada 4: 1.00%	
Chamada 5: 0.00%	
Resultado final: Chamada 1	

Exemplo de Interface interativa onde são colocadas as informações dos candidatos e é obtida a previsão de convocação.

6 Conclusão

A disponibilização da nota final do último convocado, em vez de apenas a nota de corte da chamada regular, pode fornecer uma visão mais realista para os candidatos. Isso reduz desistências desnecessárias, melhora a permanência dos interessados na lista de espera e reforça a importância de acompanhar os comunicados por e-mail da universidade.

7 Repositório com Códigos da Análise

[Link da planilha com dados](#)

[Link para o Código no Coolab](#)

8 Manual de uso de ferramenta

A ferramenta permite que o usuário insira sua **nota**, **modalidade de inscrição** e **curso desejado**, e receba como retorno as probabilidades de convocação em cada uma das chamadas (ex: chamada 0 (Regular), chamadas 1,2,3 e 4 (Listas de Espera) ou 5 (Não será Convocado), além da chamada mais provável em que seria convocado.

Para utilizar:

- **Versão terminal:** execute o script *prever_chamada()* no ambiente Python. O programa solicitará que o usuário digite a nota, a modalidade e o curso. Com base nesses dados, o sistema exibe as chances de convocação em cada chamada e informa o resultado mais provável.