



Evasão na UFPR: Identificação e Antecipação

Do perfil do evadido à construção de um modelo de previsão

Leonardo Melo
Lineu Alberto Cavazani de Freitas

Coordenadoria de Estatística e Ciência de Dados

Setembro de 2025

Sumário

1	Introdução	9
2	Análises Iniciais	11
2.1	Perfil do evadido	12
2.1.1	Formas de evasão	13
2.1.2	Evasão em função de sexo	14
2.1.3	Evasão em função de etnia	14
2.1.4	Evasão em função de estado civil	15
2.1.5	Evasão em função de turno	15
2.1.6	Evasão em função de Setor	16
2.1.7	Evasão em função de Grau Acadêmico	17
2.1.8	Evasão em função do IRA	18
2.1.9	Evasão em função de Proporção de Reprovações	19
2.1.10	Considerações sobre as análises e variáveis	21
2.2	Reprovação por disciplina	22
2.3	Evasão por curso	26
3	Modelo para previsão da evasão na UFPR	29
3.1	Regressão Logística e medidas de qualidade preditiva	31
3.1.1	Ajuste do Modelo Preditivo	33
3.1.2	Modelo Final e Interpretação dos Coeficientes	35
3.1.3	Exemplo de Aplicação	38
4	Considerações Finais	41
A	Tabelas Completas	43
A.1	Disciplinas que mais reprovam	44
A.2	Taxas de Evasão por Curso	57

Lista de Tabelas

2.1	Formas de Evasão existentes nos Sistemas Acadêmicos	13
2.2	Resumo de Estatísticas (Cinco Números e Média) para o IRA	18
2.3	Resumo de Estatísticas (Cinco Números e Média) para a Pro- porção de Reprovações	20
2.4	Top 10 Disciplinas que mais reprovam por nota e percentual médio de reprovações	24
2.5	Top 10 Disciplinas que mais reprovam por frequência e per- centual médio de reprovações	25
2.6	Taxas de Evasão Calculadas por Turma para o Curso de Ad- ministração	27
2.7	Top 10 Cursos com Maiores Taxas de Evasão	28
3.1	Estrutura da Matriz de Confusão	32
3.2	Análise Comparativa do Desempenho dos Modelos Ajustados	34
3.3	Parâmetros Estimados (β_{Setor})	37
3.4	Parâmetros Estimados (β_{Grau})	37
3.5	Parâmetros Estimados (β_{Turno})	38
A.1	Disciplinas que mais reprovam por nota e percentuais de re- provação por ano.	44
A.1	Disciplinas que mais reprovam por nota e percentuais de re- provação por ano.	45
A.1	Disciplinas que mais reprovam por nota e percentuais de re- provação por ano.	46
A.1	Disciplinas que mais reprovam por nota e percentuais de re- provação por ano.	47
A.1	Disciplinas que mais reprovam por nota e percentuais de re- provação por ano.	48
A.1	Disciplinas que mais reprovam por nota e percentuais de re- provação por ano.	49

A.2	Disciplinas que mais reprovam por frequência e percentuais de reprovação por ano.	50
A.2	Disciplinas que mais reprovam por frequência e percentuais de reprovação por ano.	51
A.2	Disciplinas que mais reprovam por frequência e percentuais de reprovação por ano.	52
A.2	Disciplinas que mais reprovam por frequência e percentuais de reprovação por ano.	53
A.2	Disciplinas que mais reprovam por frequência e percentuais de reprovação por ano.	54
A.2	Disciplinas que mais reprovam por frequência e percentuais de reprovação por ano.	55
A.2	Disciplinas que mais reprovam por frequência e percentuais de reprovação por ano.	56
A.3	Taxas Média de Evasão por Curso e Campus	57
A.3	Taxas Média de Evasão por Curso e Campus	58
A.3	Taxas Média de Evasão por Curso e Campus	59
A.3	Taxas Média de Evasão por Curso e Campus	60
A.3	Taxas Média de Evasão por Curso e Campus	61
A.3	Taxas Média de Evasão por Curso e Campus	62
A.3	Taxas Média de Evasão por Curso e Campus	63

Lista de Figuras

2.1	Taxa de Evasão das IFES.	11
2.2	Taxa de Evasão da UFPR.	12
2.3	Perfil de Evadido x Não Evadido por Sexo	14
2.4	Perfil de Evadido x Não Evadido por Etnia	15
2.5	Perfil de Evadido x Não Evadido por Estado Civil	15
2.6	Perfil de Evadido x Não Evadido por Turno	16
2.7	Evasão por Setor	17
2.8	Evasão por grau acadêmico.	17
2.9	Boxplot para IRA.	19
2.10	Boxplot para Proporção de Reprovações	20
3.1	Exemplo 1.	38
3.2	Exemplo 2.	39
3.3	Exemplo 3.	39

Parte 1

Introdução

A evasão, definida como a interrupção da trajetória do estudante antes da obtenção do diploma, constitui um fenômeno complexo e multifatorial. Suas consequências são várias e geram perdas significativas de eficiência e investimento público para a instituição e o sistema educacional. Diante disso, a compreensão dos fatores que influenciam o abandono e uso destas informações para estimar a probabilidade de um aluno evadir tornam-se uma questão central de gestão para as universidades brasileiras para desenvolver medidas direcionadas de combate à evasão com a possibilidade de acessar um aluno com alta propensão de evadir antes da ocorrência do evento.

Este relatório foi elaborado pela Coordenadoria de Estatística e Ciência de Dados (PROPLAD/CECD) em consonância com o trabalho realizado pela Comissão de Combate à Evasão e Baixa Procura da UFPR. A comissão, criada em 2025, tem o intuito de estudar e propor ações de enfrentamento ao abandono de cursos de Nível Superior na UFPR.

Para a realização deste estudo, foram consolidados dados do período de 2009 a 2024, período que abrange a utilização dos dois sistemas acadêmicos da UFPR: o SIE (adotado em 2004) e o SIGA (implantado em 2020). Os dados brutos foram extraídos e unificados em uma base única, estruturada de forma que cada linha representa um registro discente anonimizado e cada coluna, uma característica deste aluno. A base de dados final contém 107.069 observações.

Com base nos dados, buscou-se: identificar marginalmente variáveis associadas à evasão; mapear cursos e disciplinas com as maiores taxas de evasão; e propor um modelo estatístico capaz de calcular a probabilidade de abandono dos estudantes de acordo com suas características.

O restante do relatório está estruturado da seguinte forma: a Parte 2 apresenta a análise dos perfis dos alunos, das disciplinas e das taxas de evasão dos cursos; a Parte 3 se dedica à modelagem estatística, apresentando a

metodologia utilizada para prever a evasão; e a Parte 4 dispõe sobre as considerações finais.

Parte 2

Análises Iniciais

Para avaliação do cenário nacional de evasão nas Instituições Federais de Ensino Superior (IFES) pode-se utilizar dados do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Ao calcularmos uma taxa de evasão simples (definida pela razão entre o número de estudantes desistentes e o total de alunos matriculados para cada ano), observa-se uma tendência clara de crescimento contínuo dessa proporção ao longo da série histórica, conforme detalhado na Figura 2.1.

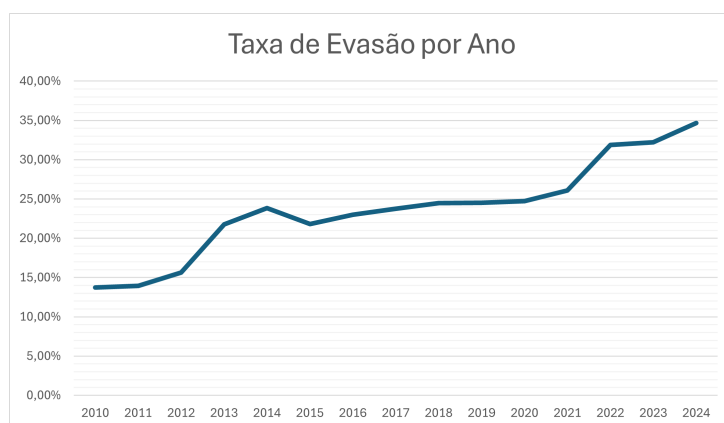


Figura 2.1: Taxa de Evasão das IFES.

Fonte: Dados do INEP.

Se calcularmos essa mesma taxa considerando dados da UFPR, conforme detalhado na Figura 2.2, verifica-se que entre 2012 e 2018

o número de evadidos manteve-se estável. Essa tendência de estabilidade foi quebrada em 2019, com um aumento considerável da evasão. Em 2020 houve uma breve queda, provavelmente influenciada pela pandemia, mas os anos seguintes registraram os picos de evasão: 2021 e 2022 atingiram os maiores níveis do período. Embora os anos de 2023 e 2024 tenham apresentado uma redução, os números de evadidos permaneceram superiores aos registrados no período de 2012 a 2018.

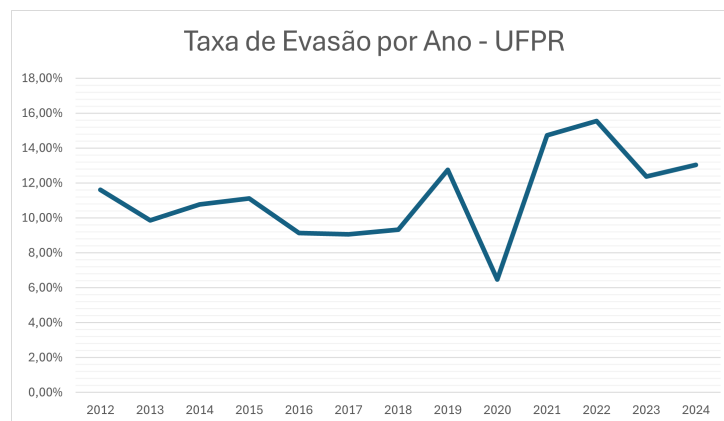


Figura 2.2: Taxa de Evasão da UFPR.

Fonte: Portal da Transparência UFPR.

O cenário delineado pela série histórica mostra que o problema é urgente e está em escalada. Este comportamento atípico justifica e fundamenta a necessidade de uma exploração profunda dos dados.

2.1 Perfil do evadido

Nesta seção são apresentados os resultados da comparação entre evadidos e não evadidos quanto a sexo, etnia, estado civil, turno, setor, Índice de Rendimento Acadêmico (IRA), proporção de reprovações e grau acadêmico.

2.1.1 Formas de evasão

A Tabela 2.1 apresenta as formas de evasão existentes nos dois sistemas acadêmicos utilizados no estudo (SIE e SIGA) e o percentual registrado para cada forma. É possível notar que a forma *Abandono* é aquela que apresenta a maior frequência no período avaliado.

Tabela 2.1: Formas de Evasão existentes nos Sistemas Acadêmicos

Forma de Evasão	Percentual (%)
Abandono	61,74
Cancelamento Pedido	7,48
Não Confirmação de Vaga	5,84
Cancelamento a Pedido	5,17
Reopção	5,00
Cancelamento a Pedido do Calouro	3,08
Novo Vestibular	2,53
Mudança de Habilitação Interna	3,09
Mudança de Turno	1,59
Desistência Vestibular	1,35
Desistência	0,60
Decisão Administrativa	0,85
Transferência Externa	0,42
Jubilamento	0,24
Falecimento	0,13
Descumprimento Edital	0,13
Desistência PROVAR	0,12
Conclusão de permanência	0,14
Reopção de Curso	0,09
Término de intercâmbio	0,10
Cancelamento Judicial	0,07
Mudança de Campus	0,08
Cancelamento por Prazo Máximo de Curso	0,05
Cancelamento Administrativo	0,07
Transferência	0,03

2.1.2 Evasão em função de sexo

A Figura 2.3 indica uma ligeira distinção no perfil por sexo. Embora o grupo de evadidos apresente uma sutil maioria de indivíduos do sexo masculino e o grupo de não evadidos mostre predominância feminina, é importante ressaltar que essa diferença se mantém em escala reduzida e não representa uma discrepância substancial.

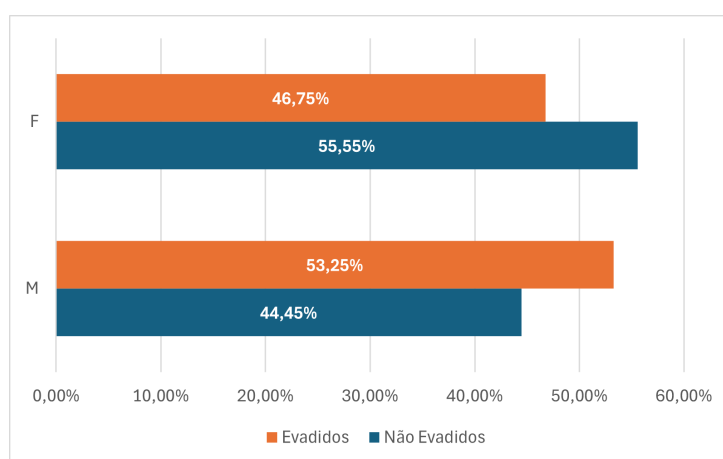


Figura 2.3: Perfil de Evadido x Não Evadido por Sexo

2.1.3 Evasão em função de etnia

A análise de evasão em função da etnia deve ser realizada com cautela por duas principais razões: a primeira delas é que não há registro de etnia em todo o período estudado; a segunda é causada pelo fato de a grande maioria dos alunos assinalarem a opção "Não desejo declarar". Para representação gráfica, a categoria "Não desejo declarar" foi retirada da análise.

Dados os devidos esclarecimentos, quando comparamos a composição das barras para evadidos e não evadidos, verifica-se que não há grandes diferenças na distribuição das categorias, conforme evidenciado pela Figura 2.4.

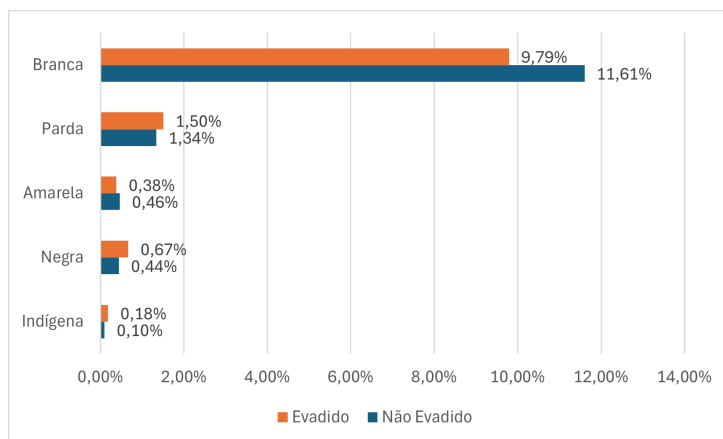


Figura 2.4: Perfil de Evadido x Não Evadido por Etnia

2.1.4 Evasão em função de estado civil

Quando avaliamos a relação entre evasão e estado civil registrado também não parece haver uma clara relação de associação entre as variáveis, como pode ser visto na Figura 2.5.

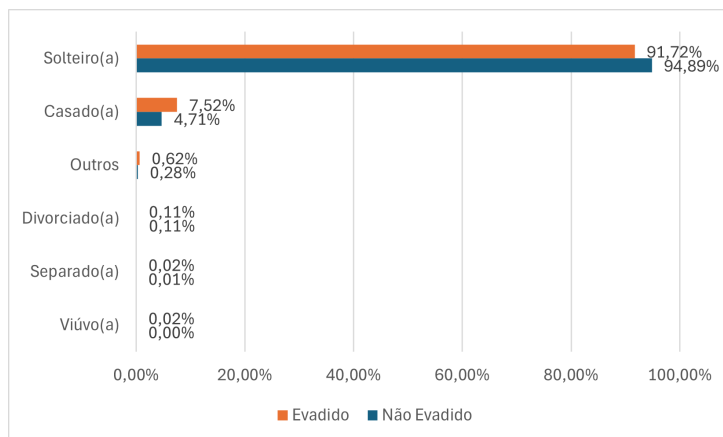


Figura 2.5: Perfil de Evadido x Não Evadido por Estado Civil

2.1.5 Evasão em função de turno

A análise de evasão em função dos turnos mostra que os não evadidos são predominantemente do turno integral, enquanto que

os evadidos são predominantemente do turno vespertino.

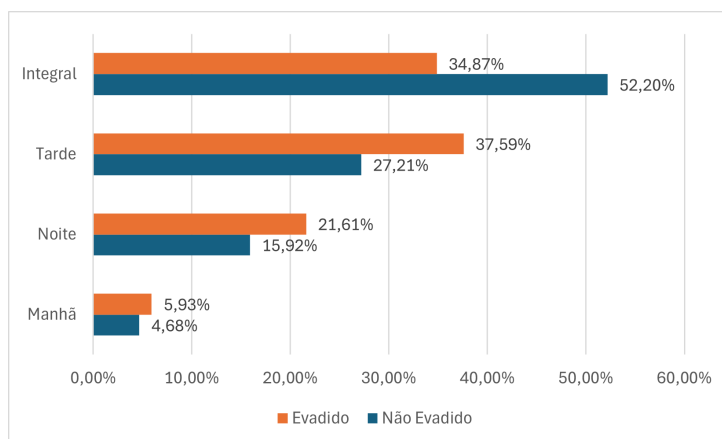


Figura 2.6: Perfil de Evadido x Não Evadido por Turno

2.1.6 Evasão em função de Setor

Dentro dos Setores da UFPR, analisamos a proporção de alunos evadidos sobre o total analisado no período. O resultado pode ser visto na Figura 2.7, onde o Setor Litoral se destaca em primeiro lugar, seguido por de Ciências Exatas e Jandaia do Sul. Já os setores com os menores índices de evasão foram os da Saúde, Jurídicas e Toledo.

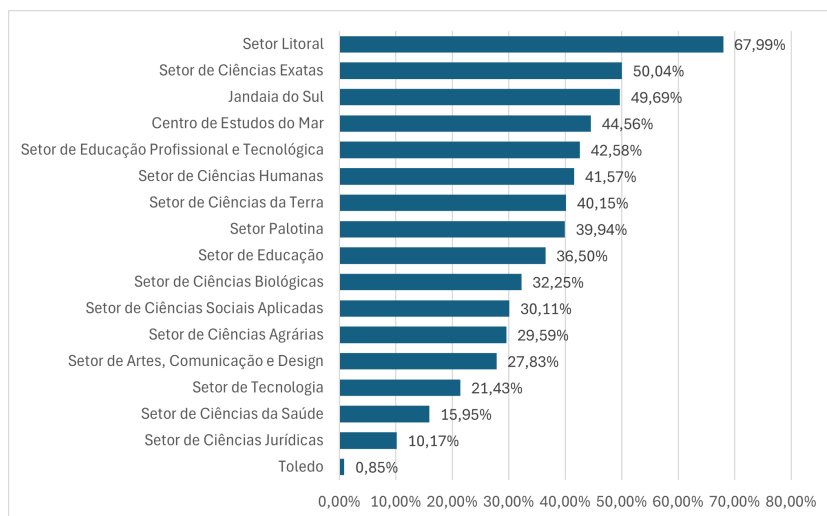


Figura 2.7: Evasão por Setor

2.1.7 Evasão em função de Grau Acadêmico

O Figura 2.8 apresenta as taxas de evasão para cada uma das quatro categorias da variável Grau Acadêmico. Podemos perceber que nos cursos ABI, a evasão é maior do que nos demais. Cursos com grau tecnológico e licenciaturas apresentam taxas próximas e chegar a trocar de posições a depender do período avaliado.

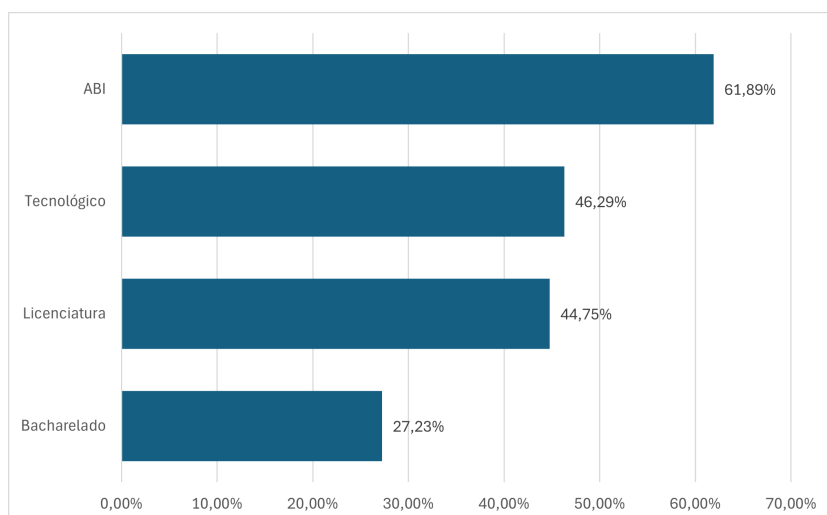


Figura 2.8: Evasão por grau acadêmico.

2.1.8 Evasão em função do IRA

O índice de rendimento acadêmico (IRA) é uma métrica a nível de aluno que varia no intervalo de 0 a 1 e é calculado por meio da seguinte expressão:

$$IRA = \frac{\sum(notas \times carga\ horaria)}{Carga\ horaria\ total} \quad (2.1)$$

Para o IRA foram calculadas medidas descritivas e o Boxplot para cada grupo. O resultado pode ser visto na Tabela 2.2 e na Figura 2.9.

Tabela 2.2: Resumo de Estatísticas (Cinco Números e Média) para o IRA

Estatística	Não evadidos	Evadidos
Mínimo (Min)	0,0773	0,0009
Primeiro Quartil (1Q)	0,7070	0,0840
Média	0,7771	0,3340
Mediana	0,7890	0,3180
Terceiro Quartil (3Q)	0,8500	0,5330
Máximo (Máx)	0,9975	0,9970

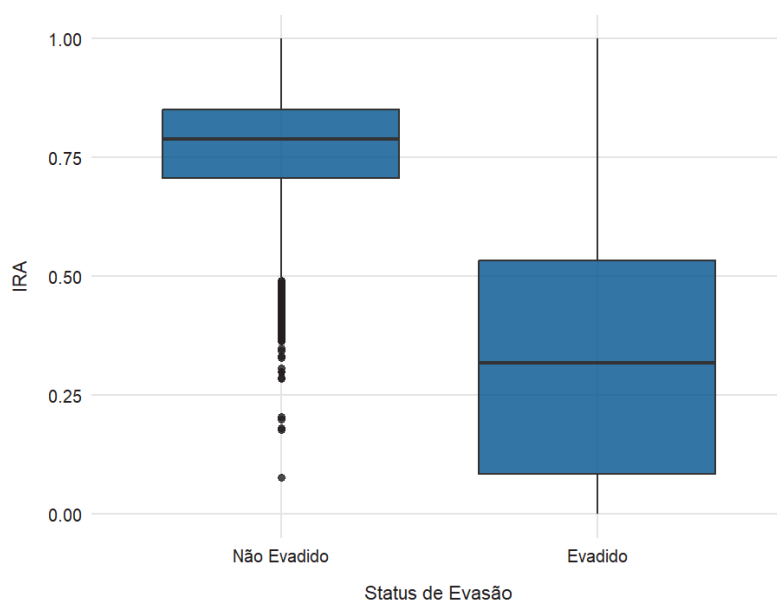


Figura 2.9: Boxplot para IRA.

Por esses resultados, percebe-se uma diferença substancial no comportamento do IRA para alunos evadidos e não evadidos, tanto na estimativa pontual da média quanto na amplitude dos dados. É possível notar que alunos não evadidos apresentam valores de IRA mais elevados do que alunos evadidos.

2.1.9 Evasão em função de Proporção de Reprovações

A Proporção de Reprovações foi calculada com base nas disciplinas cursadas e reprovadas:

$$PropRep = \frac{Disciplinas\ Reprovadas}{Disciplinas\ Cursadas} \quad (2.2)$$

De maneira análoga ao que foi feito para a variável *IRA*, foram calculados os resumos numéricos para a variável *PropRep*, cujos resultados podem ser vistos na Tabela 2.3.

Tabela 2.3: Resumo de Estatísticas (Cinco Números e Média) para a Proporção de Reprovações

Estatística	Não evadidos	Evadidos
Mínimo (Min)	0,0000	0,0000
Primeiro Quartil (1Q)	0,0000	0,0167
Média	0,0469	0,4280
Mediana	0,0154	0,3790
Terceiro Quartil (3Q)	0,0676	0,6470
Máximo (Máx)	0,6770	1,0000

Assim como no caso do IRA, a diferença entre evadidos e não evadidos é nítida de tal modo que a proporção de reprovações para evadidos é maior do que aquela observada para não evadidos.

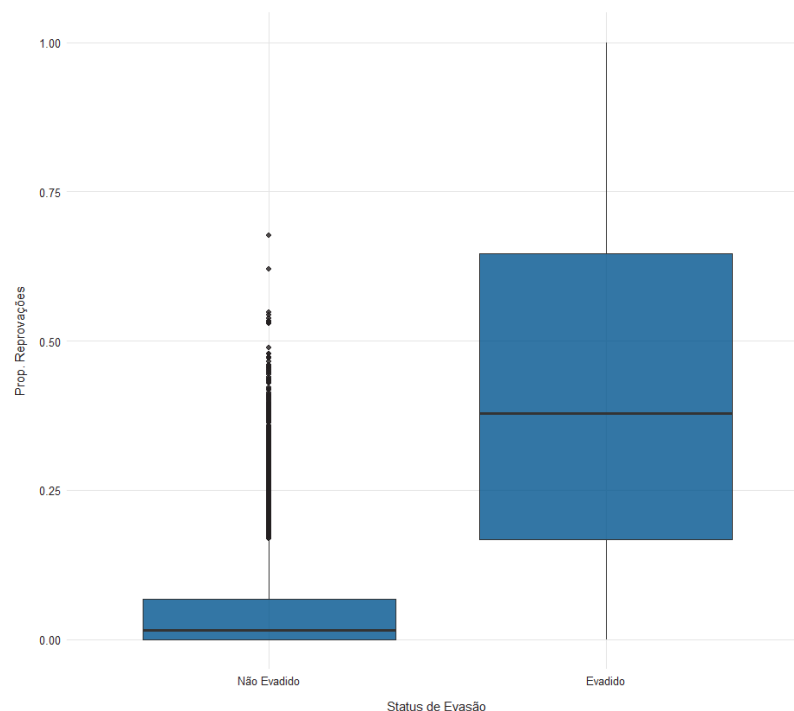


Figura 2.10: Boxplot para Proporção de Reprovações

2.1.10 Considerações sobre as análises e variáveis

As análises apresentadas evidenciam que as variáveis turno, setor, grau, IRA e proporção de reprovações são variáveis com algum potencial de discriminação entre evadidos e não evadidos.

Cabe destacar que variáveis socioeconômicas não foram incluídas na análise, por uma série de motivos. A única fonte disponível para esse tipo de informação é o questionário socioeconômico do Núcleo de Concursos (NC), aplicado quando da candidatura do aluno para prestar o Vestibular. Este questionário não foi elaborado com foco em análises longitudinais, e, por ter sido modificado ao longo dos anos, a compatibilização dos dados entre diferentes períodos tornou-se inviável. Além disso, muitas informações sensíveis são autodeclaradas e podem sofrer alteração durante a vida acadêmica do aluno, o que torna o processo de validação dos dados inviável. Por conta de todos esses fatores, e em alinhamento com o corpo técnico do NC, decidiu-se por não utilizar as variáveis do questionário socioeconômico do vestibular.

Seguindo sugestões, também explorou-se o consumo no Restaurante Universitário (RU). Foram feitas algumas análises na base oriunda de seu respectivo sistema, no período de 2017 a 2024 (o sistema foi implementado em 2016). Cruzando-se essa base com a base utilizada nas demais análises, verificamos que cerca de 55,44% dos alunos não utilizam o RU, ou seja, não traria informação válida para todos os alunos da universidade e dificultaria a criação de regras de decisão baseadas no número de idas ao RU.

2.2 Reprovação por disciplina

Com o intuito de identificar possíveis frentes de atuação a nível de cursos, verificou-se quais foram as disciplinas com maiores taxas de reprovação nos últimos sete anos. Apesar dos dados disponíveis começarem em 2009, esse período foi escolhido para minimizar o impacto de mudanças de grade curricular dos cursos, docentes que não estão mais na Universidade, entre outros motivos.

Destes últimos sete anos (2018 a 2024), foram considerados apenas casos de disciplinas que tiveram pelo menos 4 ocorrências, ou seja, que foram ofertadas em mais da metade do período avaliado. Calculou-se, então, a proporção média de reprovações por nota e por frequência nesses últimos sete anos. O resultado das 10 disciplinas com maiores índices de reprovação estão apresentados nas Tabelas 2.4 e 2.5. O resultado completo pode ser visto no Apêndice A.1.

Tabela 2.4: Top 10 Disciplinas que mais reprovam por nota e percentual médio de reprovações

NOME DA DISCIPLINA	2018 (%)	2019 (%)	2020 (%)	2021 (%)	2022 (%)	2023 (%)	2024 (%)	Média (%)
FÍSICA GERAL	0,0	43,0	0,0	26,0	24,0	36,0	36,0	33,0
FILOSOFIA MODERNA I	0,0	50,0	7,0	25,0	22,0	37,0	33,0	29,0
SOLOS I	0,0	54,0	0,0	11,0	25,0	12,0	28,0	26,0
GENÉTICA	0,0	36,0	0,0	21,0	30,0	21,0	18,0	26,0
MECÂNICA DOS SÓLIDOS	0,0	67,0	0,0	0,0	6,0	0,0	27,0	25,0
SISTEMÁTICA DE PLANTAS CULTIVADAS	0,0	50,0	0,0	8,0	17,0	11,0	35,0	24,0
PROJETOS AGRÍCOLAS	0,0	0,0	11,0	41,0	0,0	57,0	6,0	23,0
GEOMETRIA ANALÍTICA	0,0	36,0	0,0	0,0	5,0	8,0	38,0	22,0
PRÉ-CÁLCULO	0,0	0,0	0,0	3,0	12,0	47,0	22,0	21,0
TOPOGRAFIA I	0,0	43,0	0,0	0,0	0,0	8,0	28,0	20,0

Tabela 2.5: Top 10 Disciplinas que mais reprovam por frequência e percentual médio de reprovações									
NOME DA DISCIPLINA	2018 (%)	2019 (%)	2020 (%)	2021 (%)	2022 (%)	2023 (%)	2024 (%)	Média (%)	
PROBABILIDADE E ESTATÍSTICA	0,0	40,0	0,0	0,0	100,0	30,0	21,0	48,0	
QUESTÕES CONTEMPORÂNEAS DA CIÊNCIA BÁSICA	0,0	40,0	50,0	0,0	32,0	55,0	11,0	38,0	
SEGURANÇA DE SISTEMAS COMPUTACIONAIS	0,0	0,0	0,0	21,0	40,0	41,0	46,0	37,0	
PRÉ-CÁLCULO	0,0	0,0	0,0	53,0	29,0	23,0	41,0	37,0	
TECNOLOGIAS APLICADAS NA EDUCAÇÃO À DISTÂNCIA	40,0	43,0	13,0	38,0	0,0	23,0	18,0	29,0	
HISTÓRIA E FILOSOFIA DA CIÊNCIA	0,0	0,0	0,0	29,0	13,0	24,0	45,0	28,0	
PSICOLOGIA DA EDUCAÇÃO	0,0	0,0	0,0	15,0	32,0	25,0	39,0	28,0	
QUÍMICA QUÂNTICA	0,0	50,0	0,0	50,0	6,0	18,0	7,0	26,0	
GEOLOGIA PARA ENGENHEIROS	0,0	30,0	0,0	0,0	20,0	15,0	39,0	26,0	

Os resultados mostram que as disciplinas que mais reprovam por nota são em maioria da área de Ciências Exatas e da Terra, à exceção de Filosofia Moderna I. Curiosamente, existe divergência entre as disciplinas que mais reprovam por frequência, contudo as de Exatas seguem em destaque.

2.3 Evasão por curso

Calculou-se as Taxas de Evasão por Curso da seguinte forma: dentre o universo estudado, considerou-se apenas os alunos com ano de início entre 2009 e 2018, pois se trata de um período considerável para se construir uma série histórica (10 anos) e com tempo suficiente para que cada turma pudesse se formar, com poucos alunos ainda restando da turma original sem evadir ou se formar. Foram consideradas as seguintes formas de evasão:

- Abandono
- Cancelamento a Pedido
- Cancelamento por Prazo Máximo de Curso
- Decisão Administrativa
- Desistência
- Desligamento com Penalidades
- Jubilamento
- Mudança de Campus
- Mudança de Habilitação Interna
- Novo Vestibular
- Reopção

- Reopção de Curso
- Transferência
- Transferência Externa

Para cada ano de início (ou seja, para cada turma no período de 2009 a 2018) calculou-se então uma Taxa de Evasão, dada por:

$$Taxa_Curso = \frac{Total\ de\ Alunos\ Evadidos}{Total\ de\ Alunos\ da\ Turma} \quad (2.3)$$

Dessa forma, tem-se uma Taxa de Evasão por curso, por ano e foi calculada a média de taxas do período para cada curso. Como exemplo, considere as taxas calculadas para o curso de Administração, apresentadas na Tabela 2.6. A Taxa de Evasão Média para esse curso foi de 21,49%.

Tabela 2.6: Taxas de Evasão Calculadas por Turma para o Curso de Administração

Turma	Evadidos	Total Alunos	Taxa da Turma (%)
2009	49	228	21,49
2010	50	206	24,27
2011	41	214	19,16
2012	39	196	19,90
2013	55	223	24,66
2014	63	215	29,30
2015	66	263	25,10
2016	73	248	29,44
2017	87	263	33,08
2018	66	246	26,83

Seguindo essa metodologia, a Tabela 2.7 apresenta os cursos com as maiores Taxas de Evasão média para o período em questão. Os resultados completos são apresentados no Apêndice A.2.

Tabela 2.7: Top 10 Cursos com Maiores Taxas de Evasão

Nome do Curso	Taxa de Evasão Média (%)
Computação	79,82
Ciências Exatas	79,46
Informática Biomédica	78,95
Engenharia de Energia	75,23
Engenharia Agrícola	74,84
Física	74,20
História - Memória e Imagem	72,00
Engenharia Industrial Madeireira	71,64
Filosofia	70,46
Biocombustíveis	69,59

Percebe-se que os cursos das áreas de Ciências Exatas novamente aparecem em destaque.

A taxa apresentada aqui refere-se, para Computação, aos cursos de Palotina (89,11%) e Jandaia do Sul (70,53%), enquanto Ciências Exatas referem-se a Jandaia do Sul, Pontal do Paraná e Palotina (83,20%, 78,36% e 76,81%, respectivamente.)

Parte 3

Modelo para previsão da evasão na UFPR

Com o objetivo de criar uma ferramenta capaz de prever a probabilidade de evasão de um aluno, utilizamos a Regressão Logística. Esta é uma técnica estatística amplamente empregada para problemas de classificação, ou seja, para estimar a chance de um evento se concretizar.

Nesse tipo de modelo, é necessário pensar em dois elementos: variável resposta e variáveis explicativas. A variável *Resposta* é aquela que queremos prever; neste caso, se o aluno evade (1) ou não evade (0). Já as variáveis *Explicativas* são o conjunto das variáveis que podem estar associadas ao desfecho de interesse. O modelo analisa como as variáveis explicativas afetam a chance de a variável de resposta assumir o valor (1).

O procedimento consiste em treinar o modelo com dados históricos, fornecendo características dos indivíduos e o desfecho (evadiu ou não evadiu). Após o modelo ser treinado, ele é testado com um novo conjunto de dados que não foi utilizado na fase de treinamento, ou seja, o modelo preditivo é aplicado a dados aos quais ele nunca foi submetido. Essa fase de validação é crucial, pois nos permite avaliar a capacidade preditiva do modelo e garantir que ele possa ser aplicado com confiança em situações reais.

Para construção do modelo preditivo que tenha como variável

resposta a classificação do aluno entre Evadido ou Não Evadido, foram selecionadas aquelas variáveis cuja análise exploratória havia inicialmente evidenciado uma associação com a variável resposta: IRA, Proporção de Reprovações, Setor, Grau do Curso e Turno.

Conforme visto anteriormente, existem várias formas de evasão utilizadas por esses sistemas mas apenas a forma *Abandono* (a mais comum) foi utilizada e comparada apenas com alunos com a forma *Formatura*. Ou seja, alunos que não evadiram ou não se formaram durante o período analisado não foram considerados, justamente por não ter havido tempo hábil para saber se esses alunos se formaram ou evadiram.

Também há que se considerar alguns ajustes que foram realizados no intuito de melhorar a qualidade dos dados da base. Há, por exemplo, erros de cadastro nos sistemas acadêmicos, que indicam alunos com $IRA = 0$ e Forma de Evasão = Formatura. Esses casos foram excluídos da análise.

Adicionalmente, foram excluídos cursos extintos, cursos EaD e os cursos de Letras. Esses últimos não foram considerados devido à alta frequência de alterações curriculares no período analisado, o que inviabilizou a identificação de um padrão comportamental estável e robusto para ser aproveitado na modelagem preditiva.

Para mitigar o viés inerente ao desbalanceamento de classes, a base final foi submetida a um procedimento de balanceamento visando equiparar o número de observações entre as classes "Evadido" e "Não Evadido". Este processo resultou em uma base de modelagem final com um total de 52.994 observações.

3.1 Regressão Logística e medidas de qualidade preditiva

A Regressão Logística (RL) é um modelo estatístico, pertencente à família dos Modelos Lineares Generalizados (GLM), que é especificamente empregado em problemas de classificação binária. Seu objetivo é ajustar um modelo capaz de estimar a probabilidade de a variável resposta, que é discreta e assume apenas dois valores (tipicamente 0 ou 1), pertencer à classe de interesse.

O modelo chega a essa estimativa utilizando uma função de ligação (a função *logito*), que transforma a combinação linear das variáveis explicativas ($\mathbf{x}$) em uma probabilidade limitada ao intervalo $[0, 1]$. A formulação matemática relaciona o logaritmo natural da razão de chances (*odds*), $\frac{P}{1-P}$, com as preditoras:

$$\ln \left(\frac{P}{1-P} \right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

Os coeficientes de regressão ($\beta_0, \beta_1, \dots, \beta_n$) são os parâmetros do modelo e são tipicamente estimados através do método de Máxima Verossimilhança (*Maximum Likelihood Estimation* - *MLE*).

A principal vantagem da Regressão Logística é a sua interpretabilidade. Os coeficientes (β_i) representam o impacto de cada variável preditora (x_i) na razão de chances (*odds*) da variável resposta. Essa transparência permite avaliar a contribuição de cada fator para a probabilidade de ocorrência do evento.

Para realizar a classificação final, a probabilidade predita ($\hat{P}$) é comparada a um limiar de decisão (ou ponto de corte), frequentemente estabelecido em 0,50 (como foi o caso no presente estudo). Observações cuja $\hat{P}$ exceda esse limiar são classificadas como pertencentes à classe de interesse.

Por exemplo: um aluno cuja probabilidade de evadir tenha sido calculada como 0,54 seria classificado como *Evadido*, enquanto outro aluno com probabilidade de 0,48 seria classificado como *Não*

Evadido. Esse ponto de corte pode ser alterado de acordo com a realidade prática de cada situação, mas no presente estudo ele foi fixado em 0,50, mesmo porque, ao se tentar calcular um ponto de corte ótimo (que maximizasse as métricas de interesse, como Acurácia e outras), esse ponto ficou bastante próximo de 0,50.

Diferentes configurações do modelo foram ajustadas e validadas rigorosamente em dados de teste, buscando o melhor equilíbrio entre a parcimônia do modelo e sua capacidade preditiva. Para modelos de classificação binária, as Métricas de Avaliação de Desempenho são tradicionalmente derivadas da Matriz de Confusão, apresentada na Tabela 3.1. Essa matriz sumariza os quatro resultados possíveis da classificação de um modelo, comparando as classes preditas com as classes reais (observadas).

Classe Real / Predição	Predito como Classe 0	Predito como Classe 1
Classe 0	Verdadeiro Negativo (VN)	Falso Positivo (FP)
Classe 1	Falso Negativo (FN)	Verdadeiro Positivo (VP)

Tabela 3.1: Estrutura da Matriz de Confusão

Cada métrica extraída dessa matriz oferece uma perspectiva única sobre a eficácia do modelo. As medidas empregadas no presente estudo, selecionadas por sua relevância para o contexto de evasão, são detalhadas a seguir:

- **Acurácia (*Accuracy*, **ACC**)**: Representa a proporção total de predições corretas do modelo em relação ao número total de observações.

$$ACC = \frac{VP + VN}{VP + VN + FP + FN}$$

- **Sensibilidade (*Recall* ou Taxa de Verdadeiros Positivos, **SENS**)**: Mede a capacidade do modelo de identificar corretamente todas as instâncias positivas (Classe 1). No

contexto de evasão, indica quantos alunos que realmente evadiriam foram corretamente previstos.

$$SENS = \frac{VP}{VP + FN}$$

- **Precisão (*Precision*, **PREC**)**: Mede a proporção de acertos positivos que foram, de fato, corretos. De todas as observações classificadas como positivas pelo modelo, qual a fração que realmente pertence à Classe 1.

$$PREC = \frac{VP}{VP + FP}$$

- **Coeficiente de Correlação de Matthews (*Matthews Correlation Coefficient*, **MCC**)**: Uma métrica robusta e balanceada, particularmente útil em bases de dados desbalanceadas, pois considera a proporção de todos os quatro componentes da Matriz de Confusão. Seu valor varia de -1 (discordância total) a $+1$ (predição perfeita), sendo 0 um desempenho equivalente ao acaso.

$$MCC = \frac{(VP \cdot VN) - (FP \cdot FN)}{\sqrt{(VP + FP)(VP + FN)(VN + FP)(VN + FN)}}$$

Embora existam outras métricas de avaliação, estas quatro fornecem uma visão abrangente e contextualizada do desempenho preditivo do modelo, sendo suficientes para os objetivos desta pesquisa.

3.1.1 Ajuste do Modelo Preditivo

Cinco variáveis explicativas foram selecionadas para o modelo: turno, setor, grau, IRA e proporção de reprovações. Para a construção e o treinamento do modelo, foram definidos dois conjuntos de dados distintos:

1. **Base 1:** Englobando todos os *status* de desfecho (Evasão, Formatura e Sem Evasão), com o objetivo de discriminar o aluno evasor dos demais.
2. **Base 2:** Focada apenas nos desfechos Evasão e Formatura, removendo os casos com *status* Sem Evasão para uma análise binária mais pura da trajetória.

Em ambas as bases, o desequilíbrio natural entre as classes (Evadidos x Não Evadidos) exigiu a aplicação de uma técnica de balanceamento dos dados, garantindo paridade no número de observações entre as classes para evitar o viés do classificador. Após o balanceamento, os dados foram particionados em 70% para treinamento e 30% para validação, método esse que é chamado de *holdout*.

O desempenho dos modelos preditivos, após o treinamento e validação, demonstrou uma boa capacidade de generalização para o contexto de predição de evasão, conforme sumarizado na Tabela 3.2.

Tabela 3.2: Análise Comparativa do Desempenho dos Modelos Ajustados

Métrica de Desempenho	Base 1	Base 2
Acurácia (ACC)	85,31	87,72
Sensibilidade (SENS)	81,47	83,52
Precisão (PREC)	88,26	91,17
Coef. de Correlação de Matthews (MCC)	70,85	75,70

As métricas indicam que o modelo ajustado na Base 2 apresentou um desempenho marginalmente superior em todas as avaliações, muito provavelmente causado pelo fato de que os *status* "Sem Evasão" acaba não "contaminando" o aprendizado do modelo.

3.1.2 Modelo Final e Interpretação dos Coeficientes

A Probabilidade de Evasão é calculada através da função logística, que transforma a contribuição linear das variáveis preditoras em uma probabilidade limitada ao intervalo $[0, 1]$. O modelo final estimado é dado por:

$$P.E. = \frac{1}{1 + \exp(-\delta)} \quad (3.1)$$

$$\delta = 4,89 - (9,04 \cdot IRA) + (3,94 \cdot PropRep) + (\beta_{Setor}) + (\beta_{Grau}) + (\beta_{Turno}) \quad (3.2)$$

Em que:

- P.E. é a Probabilidade Estimada de Evasão.
- IRA e PropRep são variáveis contínuas (Índice de Rendimento Acadêmico e Proporção de Reprovações).
- β_{Setor} , β_{Grau} e β_{Turno} são os parâmetros estimados para as variáveis categóricas, detalhados nas Tabelas 3.3 a 3.5.

O sinal de cada parâmetro na equação indica o sentido da associação com a evasão, o que é fundamental para as ações de gestão. Um coeficiente negativo (−) indica uma associação negativa, ou seja, está associado ao menor risco de evasão. Já um coeficiente positivo (+) indica uma associação positiva, ou seja, está associado ao maior risco de evasão.

Considerando as variáveis IRA e Proporção de Reprovações, interpreta-se:

- **Índice de Rendimento Acadêmico (IRA):** O coeficiente de -9,04 é negativo, significando que quanto maior o IRA do aluno, menor é a sua probabilidade de evasão.
- **Proporção de Reprovações (PropRep):** O coeficiente de +3,94 é positivo, indicando que quanto maior a Proporção de Reprovações, maior é a probabilidade de evasão.

Para interpretação das variáveis categóricas utiliza-se categorias de referência para comparar se as demais categorias aumentam ou diminuem o risco de evasão. As categorias de referência são arbitrárias e não afetam a qualidade do modelo. Neste modelo, as classes de referência para Setor, Grau e Turno foram o Setor de Ciências da Saúde, Bacharelado e Turno Integral, respectivamente.

Para Setor:

- **Associação Negativa (Menor Risco):** Setores com coeficientes β_{Setor} negativos, como Toledo, Tecnologia e Jurídicas, apresentam menor risco associado à evasão em comparação com o Setor de Ciências da Saúde.
- **Associação Positiva (Maior Risco):** Setores com coeficientes β_{Setor} positivos, como Litoral, SACOD e Biológicas, demonstram maior risco associado à evasão em comparação com o Setor de Ciências da Saúde.

Tabela 3.3: Parâmetros Estimados (β_{Setor})

(Referência: Ciências da Saúde)

Setor	β_{Setor}	Associação
Toledo	-2,6061	← Menor Risco Associado
Jurídicas	-1,0715	← Menor Risco Associado
Tecnologia	-1,0046	← Menor Risco Associado
Agrárias	-0,6976	← Menor Risco Associado
Palotina	-0,5663	← Menor Risco Associado
Sociais Aplicadas	-0,4603	← Menor Risco Associado
Jandaia	-0,4511	← Menor Risco Associado
Exatas	-0,2844	← Menor Risco Associado
Educação	-0,2140	← Menor Risco Associado
SEPT	-0,1657	← Menor Risco Associado
da Terra	0,1050	→ Maior Risco Associado
Humanas	0,4423	→ Maior Risco Associado
CEM	0,5010	→ Maior Risco Associado
Biológicas	0,5935	→ Maior Risco Associado
SACOD	0,8110	→ Maior Risco Associado
Litoral	1,0219	→ Maior Risco Associado

Para Grau, todos apresentaram Associação Positiva, ou seja, os cursos de ABI, Licenciatura e Tecnológico apresentam coeficientes positivos, indicando que alunos nesses graus têm maior risco associado à evasão em comparação com o grau de Bacharelado.

Tabela 3.4: Parâmetros Estimados (β_{Grau})

(Referência: Bacharelado)

Grau	β_{Grau}	Associação
ABI	0,6852	→ Maior Risco Associado
Licenciatura	0,6464	→ Maior Risco Associado
Tecnológico	0,3225	→ Maior Risco Associado

Já para Turno, alunos de cursos matutinos e noturnos possuem maior risco associado, em comparação com o turno integral, enquanto que alunos de cursos vespertinos possuem menor risco associado em comparação com integral.

Tabela 3.5: Parâmetros Estimados (β_{Turno})
(Referência: Integral)

Grau	β_{Turno}	Associação
M	0,2275	→ Maior Risco Associado
N	0,2219	→ Maior Risco Associado
T	-0,1070	← Menor Risco Associado

3.1.3 Exemplo de Aplicação

A partir do modelo ajustado, é possível fazer uso da expressão para predição de casos novos não usados no treinamento e teste do modelo. A título de ilustração, imaginemos três exemplos distintos, apresentados nas figuras 3.1 a 3.3.

PARÂMETROS:			
IRA:	0,70	PROABILIDADE DE EVASÃO:	13,15%
Proporção de Reprovações:	0,10		
Setor:	Jurídicas		
Grau:	Bacharelado		
Turno	Manhã		

INTERVALO DE CONFIANÇA (95%)	
LIMITE INFERIOR:	10,50%
LIMITE SUPERIOR:	16,34%

Figura 3.1: Exemplo 1.

PARÂMETROS:		PROABILIDADE DE EVASÃO: 54,77%	
IRA:	0,60	INTERVALO DE CONFIANÇA (95%)	
Proporção de Reprovações:	0,20		
Setor:	Exatas		
Grau:	Bacharelado	LIMITE INFERIOR:	51,39%
Turno	Noite	LIMITE SUPERIOR:	58,12%

Figura 3.2: Exemplo 2.

PARÂMETROS:		PROABILIDADE DE EVASÃO: 86,85%	
IRA:	0,60	INTERVALO DE CONFIANÇA (95%)	
Proporção de Reprovações:	0,30		
Setor:	Biológicas		
Grau:	Licenciatura	LIMITE INFERIOR:	84,94%
Turno	Integral	LIMITE SUPERIOR:	88,55%

Figura 3.3: Exemplo 3.

No primeiro exemplo, temos um aluno com alto IRA, baixa proporção de reprovações, do Setor de Ciências Jurídicas, cursando um curso de Bacharelado em turno matutino. Segundo o modelo, este perfil é compatível com uma baixa probabilidade de evasão (13,15%).

No segundo caso, alteramos os parâmetros para um aluno de desempenho mediano no setor de Ciências Exatas, o que resultou em uma probabilidade de evasão moderada (54,77%).

Por fim, simulamos um caso extremo, com IRA igual ao do segundo caso, mas com maior taxa de reprovações e cursando Licenciatura no Setor de Ciências Biológicas, o que resultou em uma probabilidade de evasão elevada (86,85%).

Com a implantação do modelo em um sistema acadêmico esses cálculos podem ser realizados em tempo real de tal modo que as coordenações de curso terão em mãos uma ferramenta poderosa

para avaliar as probabilidades de evasão de cada aluno, de maneira simples e rápida e assim planejar ações de combate à evasão.

Parte 4

Considerações Finais

O presente relatório conclui um estudo abrangente sobre a evasão na UFPR contribuindo com a identificação dos principais fatores de risco e o desenvolvimento de um modelo de Aprendizado de Máquina como ferramenta de apoio à intervenção a alunos com alta propensão de evasão.

Por meio da análise exploratória dos dados disponíveis identificou-se que, marginalmente, as variáveis que possuem associação com evasão são: turno, setor, grau, IRA e proporção de reprovações.

Com as variáveis selecionadas na análise exploratória, construiu-se um modelo preditivo de Regressão Logística que alcançou valores satisfatórios nas principais métricas de qualidade preditiva, indicando um bom poder de generalização para dados não vistos.

Para garantir a escolha metodológica, este modelo foi ainda comparado com outras técnicas de Aprendizado de Máquina (como KNN, Árvore Aleatória, Floresta Aleatória, SGB, SVM e Redes Neurais). A análise comparativa demonstrou que as métricas de desempenho para todas as técnicas foram bastante similares, justificando a seleção da Regressão Logística pela sua intrínseca interpretabilidade, clareza na comunicação dos resultados e facilidade de implementação.

O principal produto deste estudo é a capacidade de antecipação. O modelo transforma a análise histórica em uma ferramenta

de intervenção, permitindo que a Universidade passe de uma atuação reativa para uma atuação preditiva e direcionada.

A aplicação prática do modelo sugere uma integração direta com os sistemas acadêmicos da UFPR. As vantagens são claras:

- **Intervenção Focada:** O modelo permite a identificação de alunos com alta Probabilidade de Evasão (P.E.), priorizando recursos de apoio (psicológico, pedagógico, financeiro) para quem mais precisa.
- **Transparência Gerencial:** A Regressão Logística oferece total transparência (interpretabilidade) sobre o motivo pelo qual o aluno está em risco, permitindo que Coordenadores de Curso definam ações específicas.
- **Fácil Integração:** A fórmula final é simples e linear, minimizando a carga computacional para implementação.

Sugere-se como encaminhamento que o resultado do modelo seja incorporado ao perfil adequado nos sistemas, de preferência em uma visualização direta para as Coordenações de Curso de tal modo que cada Coordenação poderá, assim, estabelecer sua própria linha de corte na probabilidade informada para definir um plano de ação, tornando o combate à evasão uma política de gestão contínua e baseada em dados na UFPR.

Reforça-se que a ferramenta desenvolvida tem um papel de apoio estratégico, e não de determinação absoluta. Seu uso deve ser pautado pela parcimônia, integrando-se, de forma complementar, aos diversos instrumentos de gestão e acolhimento existentes na UFPR, consolidando uma frente unificada de combate ao abandono discente.

Apêndice A

Tabelas Completas

A.1 Disciplinas que mais reprovam

Tabela A.1: Disciplinas que mais reprovam por nota e percentuais de reprovação por ano.

NOME DA DISCIPLINA	2018 (%)	2019 (%)	2020 (%)	2021 (%)	2022 (%)	2023 (%)	2024 (%)	Média (%)
Física Geral	0	43	0	26	24	36	36	33,00
Filosofia Moderna I	0	50	7	25	22	37	33	29,00
Solos I	0	54	0	11	25	12	28	26,00
Genética	0	36	0	21	30	21	18	25,20
Mecânica Dos Sólidos	0	67	0	0	6	0	27	25,00
Sistemática De Plantas Cultivadas	0	50	0	8	17	11	35	24,20
Projetos Agrícolas	0	0	11	41	0	57	6	23,00
Geometria Analítica	0	36	0	0	5	8	38	21,75
Trabalho De Conclusão De Curso Ii - Direito Público	100	33	1	4	2	6	1	21,00
Pré-Cálculo	0	0	0	3	12	47	22	21,00
Topografia I	0	43	0	0	0	8	28	19,75

Tabela A.1: Disciplinas que mais reprovam por nota e percentuais de reprovação por ano.

NOME DA DISCIPLINA	2018 (%)	2019 (%)	2020 (%)	2021 (%)	2022 (%)	2023 (%)	2024 (%)	Média (%)
Física Básica Geral 1	0	42	7	34	23	2	8	19,33
Monografia Ii	0	67	0	0	5	3	2	19,25
Trabalho De Conclusão De Curso Ii - Direito Civil	100	0	4	0	2	4	5	19,17
Estruturas Metálicas E Em Madeira	14	12	0	6	0	32	50	19,00
Fundamentos Da Economia	0	38	0	10	4	16	27	19,00
Cálculo Aplicado À Economia I	0	42	0	0	0	14	19	18,75
Bioquímica	0	20	0	0	9	20	44	18,60
Programação De Computadores	68	17	0	6	13	9	17	18,57
Cálculo I	0	25	3	5	20	24	34	18,50

Tabela A.1: Disciplinas que mais reprovam por nota e percentuais de reprovação por ano.

NOME DA DISCIPLINA	2018 (%)	2019 (%)	2020 (%)	2021 (%)	2022 (%)	2023 (%)	2024 (%)	Média (%)
Ecologia De Agroecosistemas E Sustentabilidade	0	14	0	6	15	33	23	18,20
Física Experimental I	0	30	0	0	14	11	16	17,75
Qualidade De Software	0	100	0	3	0	2	1	17,67
Geometria Analítica	0	33	0	11	15	22	23	17,33
Cálculo Diferencial E Integral Iii	0	20	0	0	36	3	9	17,00
Química Geral	0	24	4	35	7	17	14	16,83
Química Orgânica Ii	0	60	0	0	5	10	5	16,00
Álgebra Linear	13	7	4	14	23	21	29	15,86
Física Iv Para Ee	0	18	3	33	23	0	1	15,60
Estética Da Música	0	25	6	0	10	5	32	15,60
Hidrologia E Manejo De Bacias Hidrográficas	0	35	0	0	1	23	19	15,60
Álgebra Linear	0	13	25	24	0	8	23	15,50
Topografia	0	31	0	6	10	9	20	15,20

Tabela A.1: Disciplinas que mais reprovam por nota e percentuais de reprovação por ano.

NOME DA DISCIPLINA	2018 (%)	2019 (%)	2020 (%)	2021 (%)	2022 (%)	2023 (%)	2024 (%)	Média (%)
Trabalho De Conclusão De Curso Ii - Direito Penal E Processual	100	0	0	0	0	3	3	15,14
História Da Música Brasileira Ii	0	40	0	0	0	11	24	15,00
Lógica I	0	0	0	0	56	9	10	15,00
Calculo Diferencial E Integral I	19	13	0	10	31	20	7	14,29
Física F2	2	21	5	36	0	25	9	14,00
Operações Unitárias	18	4	7	0	30	12	24	13,57
Química Ii	0	11	0	0	30	18	33	13,14
Cálculo 1A	15	15	0	6	11	21	23	13,00
Prática Jurídica Trabalhista	50	20	3	8	0	3	3	12,43
Fisica Iii	12	20	0	3	10	21	20	12,29
Fisica F	17	17	0	17	8	25	0	12,00
Geometria Analítica	10	11	1	16	15	20	6	11,29

Tabela A.1: Disciplinas que mais reprovam por nota e percentuais de reprovação por ano.

NOME DA DISCIPLINA	2018 (%)	2019 (%)	2020 (%)	2021 (%)	2022 (%)	2023 (%)	2024 (%)	Média (%)
Cálculo 2A	6	8	1	7	17	10	17	9,43
Ações Constitucionais	25	5	1	0	13	7	5	9,33
Trabalho De Conclusão De Curso Ii - Direito Privado	0	25	18	6	3	7	0	8,43
Fundamentos De Matemática	26	9	0	0	2	8	5	7,14
Utilidades E Instrumentacao	19	0	0	2	7	0	10	5,43
Petrologia Sedimentar	0	0	0	0	8	20	3	5,17
Quimica Analitica - Zootecnia -	8	3	3	0	3	0	15	4,57
Física Vi	0	0	10	3	7	0	8	4,00
Decisao De Financiamento	12	2	0	0	0	5	7	3,71
Economia Florestal Ii	3	3	0	4	0	3	11	3,43

Tabela A.1: Disciplinas que mais reprovam por nota e percentuais de reprovação por ano.

NOME DA DISCIPLINA	2018 (%)	2019 (%)	2020 (%)	2021 (%)	2022 (%)	2023 (%)	2024 (%)	Média (%)
Projetos Da Indústria Química I	7	1	0	0	0	1	0	1,29

Tabela A.2: Disciplinas que mais reprovam por frequência e percentuais de reprovação por ano.

NOME DA DISCIPLINA	2018 (%)	2019 (%)	2020 (%)	2021 (%)	2022 (%)	2023 (%)	2024 (%)	Média (%)
Probabilidade E Estatística	0	40	0	0	100	30	21	47,75
Questões Contemporâneas Da Ciência Básica	0	40	50	0	32	55	11	37,60
Segurança De Sistemas Computacionais	0	0	0	21	40	41	46	37,00
Pré-Cálculo	0	0	0	53	29	23	41	36,50
Cultura E Ensino De Línguas Estrangeiras Modernas Na Escola	0	100	0	0	15	0	16	32,75
Cálculo Aplicado À Economia I	0	8	100	0	0	11	2	30,25
Tecnologias Aplicadas Na Educação À Distância	40	43	13	38	0	23	18	29,17
História E Filosofia Da Ciência	0	0	0	29	13	24	45	27,75

Tabela A.2: Disciplinas que mais reprovam por frequência e percentuais de reprovação por ano.

NOME DA DISCIPLINA	2018 (%)	2019 (%)	2020 (%)	2021 (%)	2022 (%)	2023 (%)	2024 (%)	Média (%)
Psicologia Da Educação	0	0	0	15	32	25	39	27,75
Química Quântica	0	50	0	50	6	18	7	26,20
Geologia Para Engenheiros	0	30	0	0	20	15	39	26,00
Fundamentos De Matemática	29	28	3	29	24	31	36	25,71
Introdução À Biomecânica	0	67	0	0	0	12	22	25,25
Introdução À Pintura	0	100	0	13	2	17	13	24,17
Química Orgânica I	0	0	13	29	14	48	35	23,17
Introdução À Tridimensionalidade	0	100	0	9	3	14	13	23,17
Bases Teóricas Da Administração De Ambientes De Informação	33	22	23	0	17	0	17	22,40
Metodologia De Pesquisa Em Artes 1	0	100	3	8	9	9	5	22,33

Tabela A.2: Disciplinas que mais reprovam por frequência e percentuais de reprovação por ano.

NOME DA DISCIPLINA	2018 (%)	2019 (%)	2020 (%)	2021 (%)	2022 (%)	2023 (%)	2024 (%)	Média (%)
Corpo, Gênero E Sexualidade	0	33	0	0	25	15	12	21,25
Gestão Da Inovação E Do Conhecimento	29	33	0	22	15	0	2	20,20
Int. Curricular Obrigatorio Em Tocoginecologia	0	100	0	0	0	0	0	20,00
Cálculo Diferencial E Integral I	0	0	0	0	30	24	26	20,00
Expressão Gráfica	0	0	0	0	23	27	30	20,00
Filosofia Clássica Alemã I	0	50	0	14	19	9	27	19,83
Geometria Analítica	0	0	20	22	14	25	37	19,67
Álgebra Linear	0	50	0	0	24	23	19	19,33
Ética I	0	20	24	17	21	0	32	19,00

Tabela A.2: Disciplinas que mais reprovam por frequência e percentuais de reprovação por ano.

NOME DA DISCIPLINA	2018 (%)	2019 (%)	2020 (%)	2021 (%)	2022 (%)	2023 (%)	2024 (%)	Média (%)
Informática E Banco De Dados Florestais (Aplicativos E Software)	15	12	10	14	15	31	35	18,86
Metodologia E Práticas De Ensino De Física 1	0	18	7	25	21	16	24	18,50
Introdução À Teoria Da Literatura	0	50	0	18	12	3	9	18,40
Mecânica Dos Sólidos	0	0	0	0	33	22	18	18,25
Laboratório De Física Básica 1	0	21	18	19	26	4	19	17,83
Cálculo 3B	0	35	17	9	5	25	16	17,83
Geometria Analítica	0	0	0	0	21	33	17	17,75
História Do Brasil B	4	10	0	50	11	29	19	17,57
Física Básica Geral 2	0	14	13	23	23	5	26	17,33

Tabela A.2: Disciplinas que mais reprovam por frequência e percentuais de reprovação por ano.

NOME DA DISCIPLINA	2018 (%)	2019 (%)	2020 (%)	2021 (%)	2022 (%)	2023 (%)	2024 (%)	Média (%)
Introdução À Geometria Analítica E Álgebra Linear Para Engenharia Elétrica	0	18	0	0	15	16	20	17,25
Física Geral	0	14	0	15	22	22	13	17,20
Sintaxe	0	50	0	9	6	8	13	17,20
Contabilidade Do Agro-negócio	50	8	11	13	0	0	3	17,00
Instrumentacao Computacional Em Fisica	14	14	2	17	27	25	19	16,86
Redes De Computadores	12	13	15	22	18	17	21	16,86
Cartografia	0	13	0	0	17	23	14	16,75
Tópica Em Atividades De Extensão Ii	67	0	7	6	13	10	14	16,71
Física F2	12	21	0	0	24	20	38	16,43
Cálculo I	0	0	6	41	14	14	23	16,33

Tabela A.2: Disciplinas que mais reprovam por frequência e percentuais de reprovação por ano.

NOME DA DISCIPLINA	2018 (%)	2019 (%)	2020 (%)	2021 (%)	2022 (%)	2023 (%)	2024 (%)	Média (%)
Física I	0	0	14	20	19	15	30	16,33
Economia Comportamental	0	25	0	13	13	0	14	16,25
Filosofia Da Linguagem I	0	33	0	0	7	35	22	16,17
Fundamentos Da Matemática Elementar 1	0	0	3	33	33	19	5	15,50
Cálculo 1	0	12	8	12	35	13	28	15,43
Adaptações Fisiológicas Dos Animais Ao Ambiente	18	9	0	14	13	21	17	15,33
Programação De Computadores Para Gestão Da Informação	0	23	9	36	5	7	13	13,29
Propagacao Vegetal	6	11	0	15	16	12	19	13,17
Fisica Basica Experimental I	3	10	0	13	20	21	19	12,29

Tabela A.2: Disciplinas que mais reprovam por frequência e percentuais de reprovação por ano.

NOME DA DISCIPLINA	2018 (%)	2019 (%)	2020 (%)	2021 (%)	2022 (%)	2023 (%)	2024 (%)	Média (%)
Tópica Em Atividades De Pesquisa Ii	33	10	16	5	2	8	10	12,00
Cartografia Escolar	0	0	14	11	0	12	12	9,80
Barragens De Terra E Enrocamento	0	29	0	7	16	11	2	9,29
Matemática Iii	0	0	0	30	0	0	10	5,71
Marketing E Mercados De Produtos Florestais	0	5	3	9	6	11	3	5,29
Física V	14	0	0	3	0	11	0	4,00
Operações Unitárias	0	11	4	0	0	3	7	3,57

A.2 Taxas de Evasão por Curso

Tabela A.3: Taxas Média de Evasão por Curso e Campus

Curso	Campus	Taxa Média de Evasão
Computação	Campus Palotina	89,11
Computação	Campus Jandaia do Sul	70,53
Ciências Exatas	Campus Jandaia do Sul	83,20
Ciências Exatas	Campus Pontal do Paraná	78,36
Ciências Exatas	Campus Palotina	76,81
Informática Biomédica	Campus Centro Politécnico	78,95
Engenharia de Energia	Campus Palotina	75,23
Engenharia Agrícola	Campus Jandaia do Sul	74,84
Física	Campus Centro Politécnico	74,20
História - Memória e Imagem	Campus Centro - Reitoria	72,00
Engenharia Industrial	Campus Jardim Botânico	71,64
Filosofia	Campus Centro - Reitoria	70,46
Biocombustíveis	Campus Palotina	69,59

Tabela A.3: Taxas Média de Evasão por Curso e Campus

Curso	Campus	Taxa Média de Evasão
Engenharia de Aquicultura	Campus Palotina	69,12
Engenharia de Aquicultura	Campus Pontal do Paraná	61,92
Gestão de Turismo	Campus Litoral (Matinhos)	64,53
Luteria	SEPT	64,45
Administração Pública	Campus Litoral (Matinhos)	66,29
Administração Pública	Campus Jardim Botânico	63,39
Engenharia de Alimentos	Campus Jandaia do Sul	63,96
Biotecnologia	Campus Palotina	62,36
Ciências Sociais	Campus Centro - Reitoria	61,76
Matemática Industrial	Campus Centro Politécnico	60,91
Gestão da Informação	Campus Jardim Botânico	59,78
Geografia	Campus Litoral (Matinhos)	96,15
Geografia	Campus Centro Politécnico	51,52
Linguagem e Comunicação	Campus Litoral (Matinhos)	57,18

Tabela A.3: Taxas Média de Evasão por Curso e Campus

Curso	Campus	Taxa Média de Evasão
Ciências Biológicas	Campus Palotina	66,85
Ciências Biológicas	Campus Centro Politécnico	47,36
Matemática	Campus Centro Politécnico	56,31
Estatística	Campus Centro Politécnico	56,20
Engenharia Cartográfica e de Agrimensura	Campus Centro Politécnico	55,99
Negócios Imobiliários	SEPT	55,53
Música	Campus Batel	55,22
Ciências	Campus Litoral (Matinhos)	54,80
Gestão Imobiliária	Campus Litoral (Matinhos)	54,23
Engenharia Ambiental	Campus Centro Politécnico	53,30
Gestão Pública	SEPT	52,94
Produção Cênica	SEPT	52,00
Expressão Gráfica	Campus Centro Politécnico	50,06
Educação do Campo	Campus Litoral (Matinhos)	49,94
Gestão Ambiental	Campus Litoral (Matinhos)	49,89
Gestão e Empreendedorismo	Campus Litoral (Matinhos)	47,90

Tabela A.3: Taxas Média de Evasão por Curso e Campus

Curso	Campus	Taxa Média de Evasão
Agroecologia	Campus Litoral (Matinhos)	47,26
Artes	Campus Litoral (Matinhos)	47,15
Enfermagem	Campus Centro - Saúde	45,74
Artes Visuais	Campus Juvevê (SACOD)	45,62
Ciência da Computação	Campus Centro Politécnico	44,80
Educação Física	Campus Litoral (Matinhos)	45,31
Educação Física	Campus Centro Politécnico	43,22
Análise e Desenvolvimento de Sistemas	SEPT	43,88
Engenharia Ambiental e Sanitária	Campus Pontal do Paraná	42,22
Oceanografia	Campus Pontal do Paraná	42,03
Ciências Econômicas	Campus Jardim Botânico	41,21
Engenharia Elétrica	Campus Centro Politécnico	40,74
História	Campus Centro - Reitoria	40,13
Pedagogia	Campus Centro - Reitoria	39,18
Comunicação Institucional	SEPT	37,87

Tabela A.3: Taxas Média de Evasão por Curso e Campus

Curso	Campus	Taxa Média de Evasão
Engenharia Florestal	Campus Jardim Botânico	36,41
Gestão da Qualidade	SEPT	33,83
Engenharia de Bioprocessos e Biotecnologia	Campus Palotina	49,35
Engenharia de Bioprocessos e Biotecnologia	Campus Centro Politécnico	30,51
Fisioterapia	Campus Centro Politécnico	33,32
Fisioterapia	Campus Litoral (Matinhos)	32,69
Design de Produto	Campus Juvevê (SACOD)	31,96
Design Gráfico	Campus Juvevê (SACOD)	31,61
Engenharia Mecânica	Campus Centro Politécnico	31,30
Medicina Veterinária	Campus Palotina	42,74
Medicina Veterinária	Campus Cabral (Agrárias)	19,17
Farmácia	Campus Jardim Botânico	28,91
Ciências Contábeis	Campus Jardim Botânico	28,90
Agronomia	Campus Cabral (Agrárias)	31,26
Agronomia	Campus Palotina	21,97
Geologia	Campus Centro Politécnico	27,02

Tabela A.3: Taxas Média de Evasão por Curso e Campus

Curso	Campus	Taxa Média de Evasão
Administração	Campus Jardim Botânico	25,32
Biomedicina	Campus Centro Politécnico	25,14
Engenharia Civil	Campus Pontal do Paraná	34,08
Engenharia Civil	Campus Centro Politécnico	19,94
Nutrição	Campus Jardim Botânico	23,51
Engenharia Química	Campus Centro Politécnico	23,47
Psicologia	Campus Centro - Santos Andrade	22,88
Arquitetura e Urbanismo	Campus Centro Politécnico	22,79
Direito	Campus Centro - Santos Andrade	20,58
Engenharia de Produção	Campus Jandaia do Sul	27,52
Engenharia de Produção	Campus Centro Politécnico	11,11
Jornalismo	Campus Juvevê (SACOD)	17,69
Odontologia	Campus Jardim Botânico	14,18
Medicina	Unidade Acadêmica Toledo	2,72

Tabela A.3: Taxas Média de Evasão por Curso e Campus

Curso	Campus	Taxa Média de Evasão
Medicina	Campus Centro - Saúde	2,15