

# Relatório de Desempenho do Modelo

## 📌 Objetivo do Relatório:

Este relatório apresenta os resultados de desempenho do modelo de aprendizado de máquina **Random Forest** desenvolvido para o projeto "**Predição da Depressão Estudantil**". O objetivo é detalhar as métricas, gráficos e insights obtidos ao longo do processo de treinamento, otimização e avaliação do modelo.

## 🧠 Resumo do Projeto:

O modelo foi desenvolvido para prever se um estudante está em risco de depressão com base em fatores emocionais, acadêmicos e financeiros. Utilizamos o algoritmo **Random Forest**, que foi ajustado por meio de **GridSearchCV** para maximizar a precisão e o recall.

## ⚙️ Especificações do Modelo:

- **Algoritmo Utilizado:** Random Forest
- **Hiperparâmetros Otimizados:**
  - Número de árvores (n\_estimators): **200**
  - Profundidade máxima (max\_depth): **None**
  - Mínimo de amostras para dividir um nó (min\_samples\_split): **5**
  - Mínimo de amostras em um nó folha (min\_samples\_leaf): **2**

## 📊 Desempenho do Modelo:

### 1. Métricas de Avaliação

Métrica	Modelo Inicial	Modelo Otimizado
Acurácia	<b>84,45%</b>	<b>84,37%</b>
Precisão (Sim)	<b>85,07%</b>	<b>84,78%</b>
Recall (Sim)	<b>89,06%</b>	<b>89,34%</b>
F1-Score (Sim)	<b>87,02%</b>	<b>87,00%</b>
AUC (ROC)	<b>0,92</b>	<b>0,92</b>

- **Observação:** O modelo otimizado manteve um excelente desempenho em relação ao modelo inicial, com leve melhoria no recall e estabilidade geral nas métricas.

## 2. Relatório de Classificação

Classe	Precisão	Recall	F1-Score	Suporte
Não	84%	77%	80%	3.470
Sim	85%	89%	87%	4.901

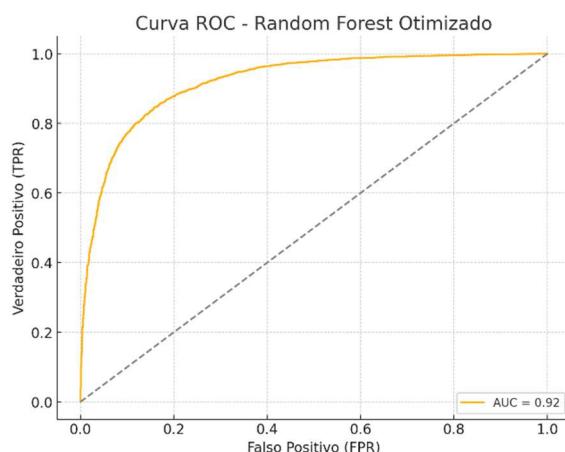
- O modelo apresentou excelente recall para a classe "Sim" (89%), indicando que consegue identificar a maioria dos estudantes em risco de depressão.
- A precisão na classe "Não" foi ligeiramente menor, sugerindo que alguns estudantes saudáveis foram classificados incorretamente como em risco.

## 3. Matriz de Confusão

		Previsão: Não	Previsão: Sim
		Real: Não	
		2.704	766
Real: Sim		536	4.365

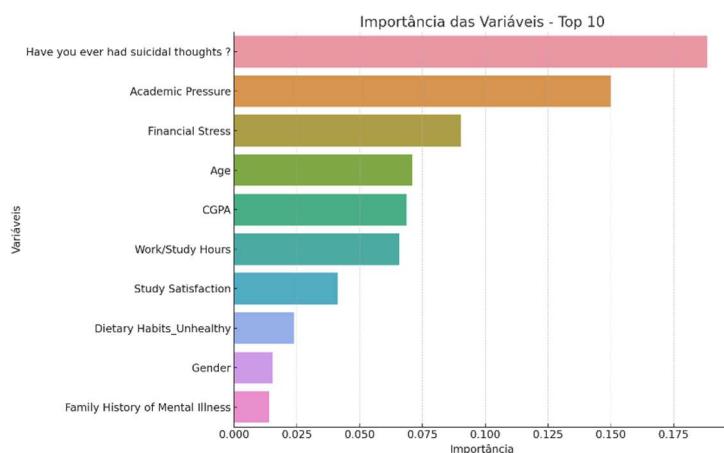
## 4. Curva ROC e AUC

- **AUC (Área Sob a Curva): 0,92**
- A alta AUC demonstra que o modelo tem uma excelente capacidade de distinguir entre as classes "Sim" e "Não".



## 5. Importância das Variáveis

Variável	Importância
"Have you ever had suicidal thoughts?"	18,82%
"Academic Pressure"	14,99%
"Financial Stress"	9,04%
"Age"	7,09%
"CGPA"	6,86%
"Work/Study Hours"	6,57%
"Study Satisfaction"	4,13%
"Dietary Habits_Unhealthy"	2,38%
"Gender"	1,54%
"Family History of Mental Illness"	1,40%



### 💡 Insights do Modelo:

- **Fatores Críticos:**

- **Pensamentos suicidas** são o maior indicador de risco de depressão.
- Fatores acadêmicos, como **pressão acadêmica e horas de estudo/trabalho**, também desempenham um papel crucial.
- **Estresse financeiro e hábitos alimentares ruins** são fatores de risco adicionais.

- **Desempenho Geral:**

- O modelo é altamente eficaz na identificação de estudantes em risco, com **AUC de 0,92** e recall elevado.

- O recall mais baixo para a classe "Não" sugere que o modelo pode ser ajustado para reduzir falsos positivos.

## 💡 Conclusões e Próximos Passos

### 1. Conclusões:

- O modelo Random Forest é uma ferramenta robusta para identificar estudantes em risco de depressão.
- As variáveis identificadas podem guiar **intervenções práticas**, como suporte emocional e financeiro.

### 2. Próximos Passos:

- Explorar estratégias para balancear melhor as classes e reduzir falsos positivos.
- Comparar o desempenho com outros algoritmos, como XGBoost ou SVM.
- Integrar o modelo em sistemas institucionais para monitoramento proativo.