

Project Overview: Predição da Depressão Estudantil

📌 Resumo do Projeto:

A depressão estudantil é um problema crescente que afeta a saúde mental, desempenho acadêmico e qualidade de vida dos estudantes. Este projeto utiliza machine learning para prever a probabilidade de um estudante estar em risco de depressão, com base em fatores como:

- Pressão acadêmica e financeira.
- Satisfação nos estudos.
- Hábitos de sono e alimentação.
- Histórico familiar de doença mental.

Nosso objetivo é fornecer insights para intervenções direcionadas, promovendo um ambiente mais saudável e suporte para estudantes em risco.

🧠 Metodologia

1. Coleta de Dados:

- Dataset: [Student Depression Dataset](#)
- Tamanho: **27.901 registros com 18 colunas.**

2. Etapas do Pipeline:

- **Análise Exploratória (EDA):**
 - Exploração de variáveis categóricas e contínuas.
 - Identificação de padrões e fatores de risco.
- **Pré-processamento:**
 - Tratamento de valores ausentes.
 - Codificação de variáveis categóricas.
 - Padronização de variáveis numéricas.

- **Construção do Modelo:**
 - Algoritmo inicial: **Random Forest**.
 - Tuning de hiperparâmetros com **GridSearchCV**.
 - **Avaliação:**
 - Métricas: Acurácia, Precisão, Recall, F1-score e AUC.
 - Visualizações: Curva ROC, Importância das Variáveis.
-

Resultados do Modelo:

1. Desempenho:

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

- O modelo otimizado manteve **alto desempenho**, com **excelente recall** para a classe "Sim" (89,3%).
- A **AUC de 0,92** confirma a capacidade robusta do modelo em distinguir entre estudantes com e sem depressão.

2. Matriz de Confusão:

		Previsão: Não	Previsão: Sim
Real: Não	Real: Não	2704	766
	Real: Sim	536	4365

3. Importância das Variáveis: As 5 variáveis mais importantes no modelo são:

- "Have you ever had suicidal thoughts?" (18,82%)
- "Academic Pressure" (14,99%)
- "Financial Stress" (9,04%)
- "Age" (7,09%)
- "CGPA" (6,86%)

4. Curva ROC:

- A alta AUC (0,92) indica que o modelo tem excelente capacidade de classificação.
-

Insights Principais:

1. Fatores Críticos:

- Estudantes que relataram **pensamentos suicidas** apresentam maior probabilidade de depressão.
- Altos níveis de **pressão acadêmica e estresse financeiro** aumentam significativamente o risco.
- **Satisfação nos estudos e hábitos alimentares** são indicadores relevantes.

2. Desempenho do Modelo:

- O modelo é robusto, com recall elevado e desempenho consistente.
 - Leve redução de falsos negativos, garantindo maior identificação de estudantes em risco.
-

Propostas de Intervenção:

Identificação Proativa:

- Usar o modelo para monitorar fatores críticos e gerar alertas para estudantes em risco.

Apoio Acadêmico:

- Reduzir pressão acadêmica com programas de tutoria e flexibilidade de prazos.

Assistência Financeira:

- Criar programas de bolsas e auxílio para estudantes em vulnerabilidade.

Saúde Mental:

- Implementar serviços de aconselhamento psicológico.
- Realizar campanhas de conscientização sobre saúde mental.

Promoção de Hábitos Saudáveis:

- Incentivar melhores rotinas de sono e alimentação.
 - Promover atividades físicas regulares.
-

Conclusão

O modelo de aprendizado de máquina desenvolvido mostrou-se eficaz na **identificação de estudantes em risco de depressão**, fornecendo insights valiosos para intervenções direcionadas. Com base nos resultados, instituições educacionais podem criar estratégias que combinem suporte emocional, acadêmico e financeiro para melhorar a saúde mental dos estudantes.