

# Inteligência Artificial

Projeto A3

# Employee DataBase

- Este conjunto de dados contém informações anônimas sobre os funcionários de uma empresa como:

Formação educacional;  
Histórico profissional;  
Dados demográficos;  
Fatores relacionados ao emprego.

Pesquisadores, analistas de dados e profissionais de RH podem obter insights importantes sobre a força de trabalho.

# Colunas

1. 'Education' - Qualificação;
2. 'JoiningYear' - Ano de início;
3. 'City' - Cidade;
4. 'PaymentTier' - Diferença salarial;
5. 'Age' - Idade;
6. 'Gender' - Gênero;
7. 'EverBenched' - Tempo desempregado;
8. 'ExperienceInCurrentDomain' - Experiência;
9. 'LeaveOrNot' - Viagens a trabalho.



## Variável 'Target'

- A variável 'Target' escolhida, 'EverBenched', tem como principal objetivo verificar se há padrões comportamentais relacionados ao afastamento de funcionários ou à falta de atribuições de tarefas.



# Edições Necessárias

Retiradas:

'City' , 'JoiningYear'

Transformações de 'String' para 'Int':

Education

Gender

EverBenched



## Primeiro Método de Aprendizagem de Máquina

Como primeiro método de aprendizagem de máquina utilizamos a **Regressão Logística** que nos deu os seguintes resultados:

accuracy			0.40	1862
macro avg	0.11	0.16	0.12	1862
weighted avg	0.26	0.40	0.29	1862

Os resultados indicam que o modelo pode precisar de melhorias, como em termos de precisão e recall para várias classes.

## Segundo Método de Aprendizagem de Máquina

Como segundo método de aprendizagem de máquina utilizamos KNN (K - *Nearest Neighbors*) que deram os seguintes resultados:

```
[[1036 15]  
 [113   0]]  
0.8900343642611683
```

Com esse método obtivemos 89% de acerto em relação a pessoas que estão empregadas e desempregadas.

## Terceiro Método de Aprendizagem de Máquina

Como terceiro método de aprendizagem de máquina utilizamos Naive Bayes que retornaram os seguintes resultados:

```
A acurácia do modelo 3 foi 90.64%  
[[1055    0]  
 [ 109    0]]
```

Com esse método obtivemos 90.64% de acerto em relação a pessoas que estão empregadas e desempregadas.





# Resumo

1

## **Regressão Logística**

Método pouco eficaz, não direcionado para esse tipo de análise.

2

## **KNN**

Método bom, podendo sofrer melhorias.  
Mostrando 89% de assertividade.

3

## **Naive Bayes**

Método mais eficaz mostrando 90.6% de acerto



T

## Conclusão

Os três testes de análise de dados revelaram insights sobre a predição do status de emprego. O modelo de regressão mostrou uma acurácia de 40%, ou seja, não é possível identificar se as pessoas já ficaram desempregadas ou não. Em contrapartida, o KNN obteve 89% de acurácia, indicando melhor desempenho na identificação de padrões. O Naive Bayes alcançou a maior acurácia, com 90.64%, destacando sua eficácia e assertividade na análise de dados relacionados ao desemprego. Os resultados enfatizam a importância de algoritmos mais robustos, como KNN e Naive Bayes, para uma predição mais precisa e significativa sobre os padrões comportamentais relacionados ao afastamento de funcionários ou à falta de atribuições de tarefas.



# Participantes

Letícia Keise RA: 822160543

Giovanni Mascotte RA: 822135846

Maria Eduarda Pereira RA: 822145642

Pedro Quintela RA: 822150807