

**Alunos: Gabriel Tomazini Marani                      RA: 2266083**  
**Paulo Victor Nogueira Rodrigues              RA: 2265125**

## **O PROBLEMA**

O problema escolhido foi prever a entrada num programa de pós graduação de uma universidade indiana através de alguns parâmetros importantes. o Dataset utilizado pode ser encontrado em:

<https://www.kaggle.com/datasets/mohansacharya/graduate-admissions>

Os parâmetros a serem considerados foram:

### **GRE Scores**

Graduate Record Examination, é um exame internacional que serve para medir o conhecimento adquirido durante a graduação, o valor da nota varia de 0 a 340.

### **TOEFL Scores**

Test of English as a Foreign Language, é um teste realizado para medir a proficiência na língua inglesa, pois a pós graduação seria conduzida em inglês, o valor da nota varia de 0 a 120.

### **University Rating**

Rank das melhores universidades, 1 sendo a pior universidade e 5 sendo a melhor.

### **Statement of Purpose and Letter of Recommendation Strength ( out of 5 )**

Declaração e objetivo na universidade e força da carta de recomendação, valor varia de 1 a 5, com 1 sendo a pior nota e 5 sendo a melhor nota.

### **Undergraduate GPA**

Grade Point Average, é a média das notas que o aluno obteve durante a graduação, o valor varia de 0 a 10.

### **Research Experience**

Experiência em pesquisa, 0 para não, 1 para sim.

### Chance of Admit

Chance de admissão, esse valor varia de 0 a 1.

## A ABORDAGEM PARA A SOLUÇÃO

Com todos esses parâmetros no banco de dados, o objetivo era prever a admissão na universidade indiana, o label escolhido portanto é “Chance of Admit”, entretanto, o valor varia entre 0 e 1, portanto, houve um pré-processamento do banco de dados de dados, onde trocamos valores menores que 50% para 0 e qualquer valor maior ou igual a 50% para 1, dessa forma agora temos um modelo que irá prever resultados de forma discreta, O método classificador escolhido foi o Random Forest.

## RESULTADOS

Após o modelo ser treinado, obtivemos uma precisão de 99% nos dois testes realizados, no primeiro teste, colocamos o melhor resultado possível, com nota máxima em todos os parâmetros e no segundo teste o contrário, com notas mínimas em todos eles.

```
Python
GREScore = 340
TOEFLScore = 120
UniversityRating = 5
SOP = 5
LOR = 5
CGPA = 10
Research = 1
```

```
[1]
[[0. 1.]]
0.996
```

Figura 1 - Entrada na universidade Indiana, 99% de acurácia, parâmetros de entrada ótimos

Python

```
GREScore = 90  
TOEFLScore = 40  
UniversityRating = 5  
SOP = 1  
LOR = 1  
CGPA = 2  
Research = 0
```

Com os parâmetros acima foi possível obter o seguinte resultado:

```
[0]  
[[0.9 0.1]]  
0.998
```

Figura 2 - Entrada na universidade Indiana, 99% de acurácia, parâmetros de entrada ruins

Como podemos observar, o algoritmo apontou para a não entrada devido as notas baixas dos parâmetros de entrada.