

Aula 3

Iyan Lucas

March 3, 2022

1 Pré-processamento

1.1 Balanceamento

Quando há muito mais atributos de uma classe do que o outro, como se tivesse 100x a mais de um atributo X do que do Y. Para saber se a base está desbalanceada, veja o recall e a precisão.

Há formas de corrigir este problema, uma destas formas é redefinir os datasets, que é retirar ou adicionar instância.¹ De forma aleatória ou induzida (heurística).

¹Undersampling ou Oversampling.