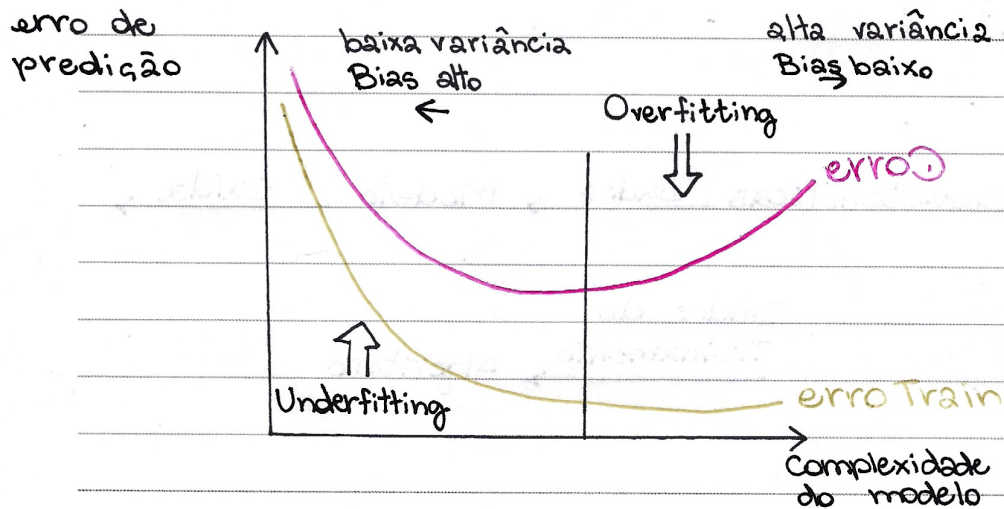


## O que é Overfitting (Sobreajuste)?

- \* Termo utilizado pela estatística.
- \* Representa o sobreajuste do modelo a um conjunto específico de dados.
- \* Leva à perda da capacidade de generalização do modelo.



### Causas do Overfitting:

- Ruídos nos dados
- Falta de Amostras de diferentes classes para treinamento
- Modelo muito complexo para resolver o problema

### Como evitar:

- Dividir os dados em treinamento (80%) e teste (20%)
- Aplicar validação cruzada (cross-validation)
- Aplicar técnicas de regularização
- Remover neurônios/camadas da rede neural (dropout)
- Aplicar a "parada precoce" (early stopping)

Minimizando o erro:  $E_{in}(w) + \frac{\lambda}{N} \cdot w^T \cdot w$  para diferente  $\lambda$



O que é Underfitting?

Quando a tarefa possui maior complexidade do que o algoritmo selecionado.

Dimensionalidade

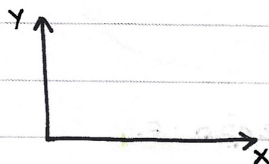
O que é dimensionalidade?

- Representa a dimensão espacial dos dados
- Diferente de ter vários datasets
- Cada dimensão perpendicular a outras

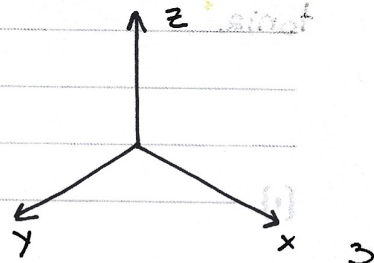
1 dimensão:



2 dimensões:



3 dimensões:

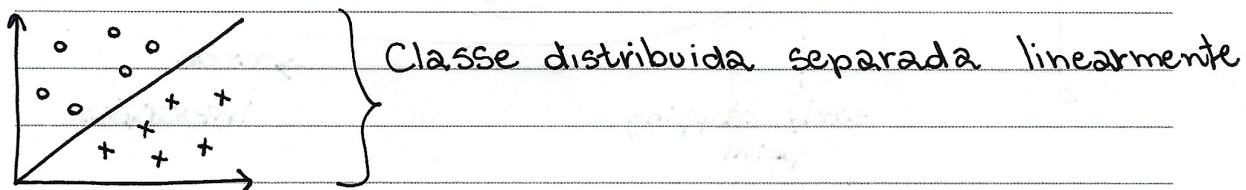




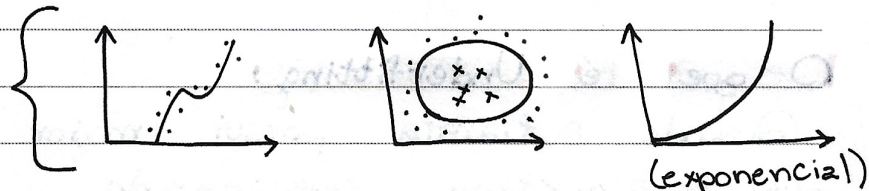
## Maldição da Dimensionalidade?

- É quando as dimensões do espaço de características aumenta.
- O número de configurações pode aumentar exponencialmente
- O número de configurações em uma observação diminui (densidade)
- Com uma menor densidade, mais instâncias são necessárias.

## Linearidade x Não Linearidade



Exemplos de Não linearidade



Passo a passo para o sucesso:

