

# Lidando com Classes Desbalanceadas

Minerando Dados.



# Machine Learning

- **Classes desbalanceadas.**
- **Impacto nos modelos.**



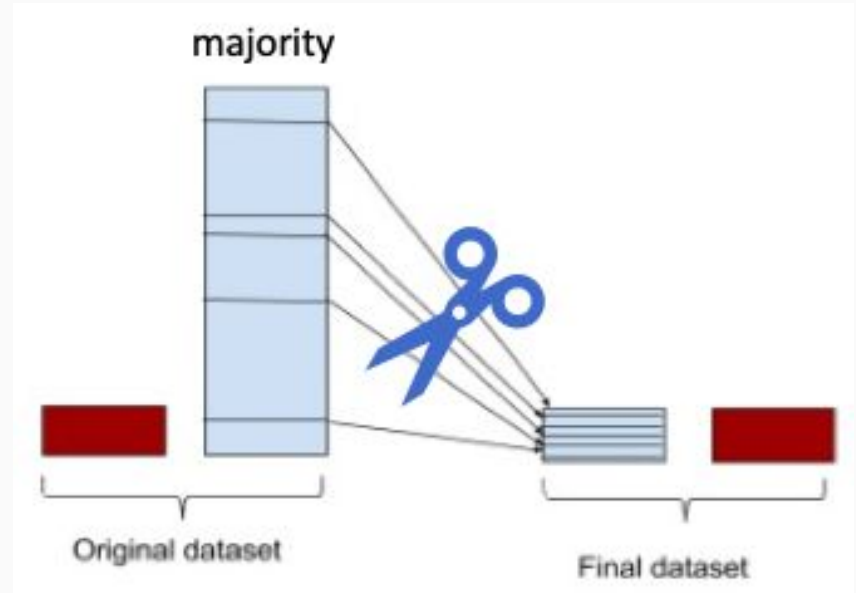
# Machine Learning

- Sistema de detecção de Fraudes.
- 90% de transações não são fraudes
- 10% de transações são fraudes.



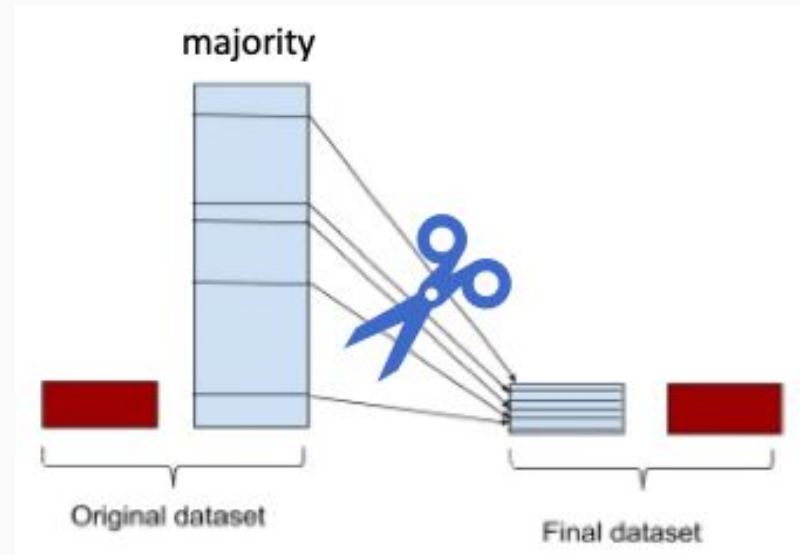
# Machine Learning

- **Undersampling**
  - Consiste em reduzir de forma aleatória os exemplos da classe majoritária.



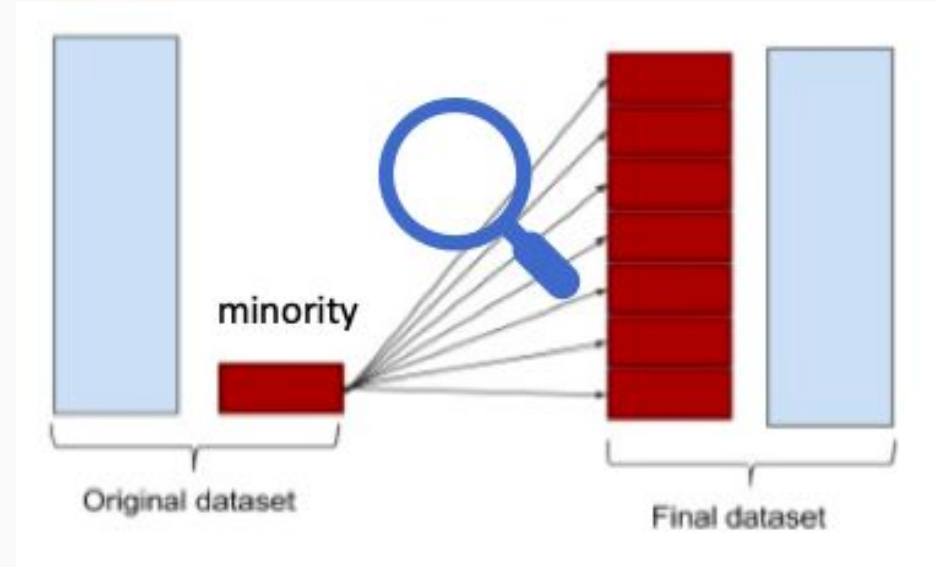
# Machine Learning

- **NearMiss**
  - Consiste em reduzir de forma aleatória os exemplos da classe majoritária.



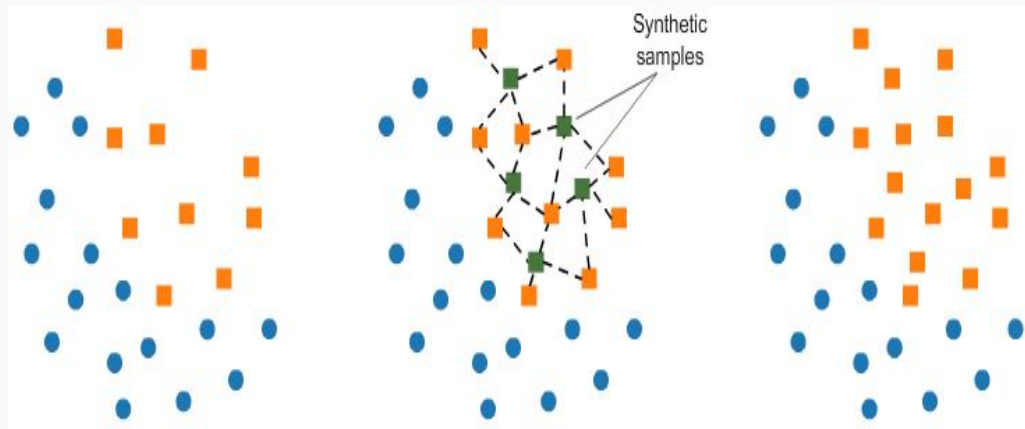
# Machine Learning

- **Oversampling**
- Consiste em replicar dados aleatórios da classe minoritária.



# Machine Learning

- **Smote**
  - Consiste em gerar dados sintéticos da classe minoritária a partir de vizinhos.



# Machine Learning

- **Engenharia de Features**
- A geração de novas features tende a dar mais robustez ao modelo e melhora a classificação da classe minoritária.





# Machine Learning

**Hands on!**