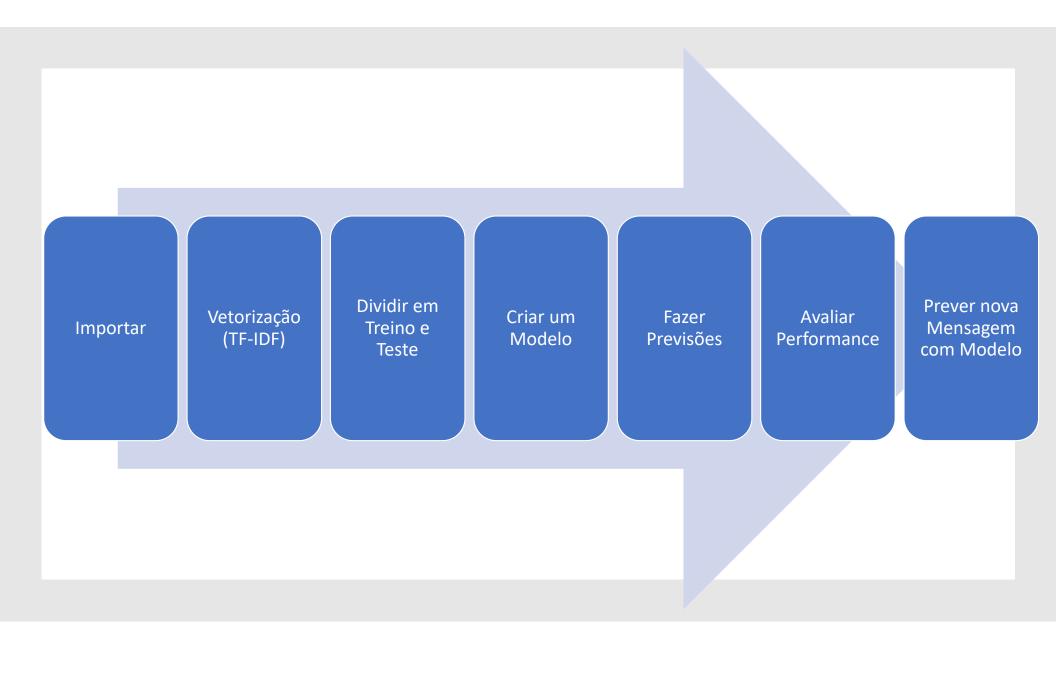
## Classificação de Spam

- Uma mensagem de e-mail é spam ou é legítima?
- Havendo dados previamente rotulados podemos treinar um modelo para prever novas mensagens

#### Dados

• 5572 mensagens

	Category	Message
0	ham	Go until jurong point, crazy Available only
1	ham	Ok lar Joking wif u oni
2	spam	Free entry in 2 a wkly comp to win FA Cup fina
3	ham	U dun say so early hor U c already then say
4	ham	Nah I don't think he goes to usf, he lives aro



### Divisão de Treino e Teste

• X\_treinamento, X\_teste, y\_treina mento, y\_teste

#### Modelo

- Random Forest ensemble
  - Cria amostras aleatórias com reposição
  - Induz n árvores de decisão aleatórias de cada amostra
  - Executa um processo de votação
  - Seleciona a previsão com mais votos

# Avaliar Performance

- Matriz de Confusão
- Acurácia
- Precision
- Recall
- F1