**Métodos preditivos**

**Classificação**

Define classes para cada registro, possui rótulos. É um algoritmo para classifica, como por exemplo:

* Satisfação do cliente;
* Risco de crédito;
* Filtros de Spam;
* Separação de notícia;
* Reconhecimento de voz/facial;
* Previsão de doenças.

**Regressão**

Similar a classificação, regressão faz previsão de valores números baseados nos dados históricos, exemplo:

* Gastos propagando -> valor de venda;
* Temperatura, umidade e pressão do ar -> velocidade do vento;
* Fatores externos -> valor do dólar;
* Resultado do exame -> probabilidade de um paciente sobreviver;
* Risco de investimento;
* Gastos no cartão de crédito, histórico -> limite

**Métodos Descritivos**

**Associação**

Como exemplo, seria o carrinho de compra, se uma pessoa compra um determinado produto, provavelmente levará um outro.

* Organização de Prateleiras de mercado (fraldas e cervejas);
* Promoções com itens vendidos em conjunto;
* Planejas catálogos das lojas e folhetos de promoções;
* Controle de evasão em universidades.

**Agrupamento**

Analisar os dados e encontrar grupos. Exemplos:

* Segmentação de mercado;
* Encontrar grupos de clientes que irão comprar um produto (mala direta);
* Agrupamento de documentos/noticias;
* Agrupamento de produtos similares;
* Perfis de clientes (Netflix);
* Análise de redes sociais.

**Detecção de desvios (outliers)**

Algo que esteja fora do padrão/fora da curva (outliers). Exemplos:

* Fraude em cartão de crédito;
* Intrusão em redes;
* Uso de energia elétrica, água ou telefone;
* Desempenho de atletas (doping);
* Monitoramento de máquinas em um data center.

**Descoberta de padrões sequenciais**

Oferta produtos sequenciais, se uma pessoa compra o primeiro livro do Harry Potter, depois de um tempo a loja envia um e-mail ofertando o segundo livro do Harry Potter. Exemplos:

* Livrarias, loja de equipamentos de atletismo, computadores;
* Marketing direcionado para pessoas que tem maiores chances de adquirir um novo produto;
* Prevenção de doenças;
* Navegação em sites

**Sumarização**

Pega um texto maior e gera um texto menor (sumarização de documentos). Exemplos:

* Perfis de pessoas -> São ouvintes do programa homens na faixa de 25 a 35 anos, com nível superior e que trabalham na área de administração;
* Segmentação de mercado.

**Tipos de Aprendizagem de Máquina**

**Supervisionada**

**Tem lições e é treinado.**

**Fase 1**: Extração de características, submeter ao Algoritmo de aprendizagem (Supervisor para classificar) e gera um Modelo aprendido.

**Fase 2**: Extração de características, submeter ao Modelo aprendido

**Não-supervisionada**

Não treina, é executado e não possui a coluna classe na tabela.

* Analisar automaticamente os dados (associação, agrupamento)
* Necessita análise para determinar o significado dos padrões encontrados

**Por reforço**

* Aprender com as interações com o ambiente (causa e efeito)
* Aprender com sua própria experiência
* Robô coletando lixo aprendendo a andar em um ambiente

**Variáveis**

**Numéricas**

* Contínua:
  + Números **reais** (float)
  + Temperatura, altura, peso, salário
* Discreta:
  + Conjunto de valores finito (**inteiros**)
  + Contagem de alguma coisa

**Categóricas**

* Nominal
  + Dados não mensuráveis
  + Sem ordenação: cor dos olhos, gênero
* Ordinal
  + Categorizado sob uma **ordenação**
  + Tamanho P, M e G

**Naive bayes**

Algoritmo de probabilidade, analisa uma base dados histórica para gerar uma outra tabela de probabilidade.

Vantagens

* Rápido;
* Simplicidade de interpretação;
* Trabalha com altas dimensões;
* Boas previsões em bases pequenas.

Desvantagem

* Combinação de características (atributos independentes) – cada par de características são independentes – nem sempre é verdade.

**Árvore de decisão**

Duas fórmulas para cálculo da árvore de decisão, analisa atributo por atributo para saber qual é o mais importante:

Entropy(S) (entropia) = -pi \* log(pi; 2) : verifica qual o melhor atributo

Gain(S, A) (Ganho de informação) = Entropy(S) - (pi \* Entropy(S))

**Random Forest (Floresta Randômica)**

Uma melhoria dos algoritmos de Árvore de decisão, utiliza a resposta de várias árvores.

* Ensemble learning (aprendizagem em conjunto)
  + “Consultar diversos profissionais para tomar uma decisão”
  + Vários algoritmos juntos para construir um algoritmo mais “forte”
  + Usa média (regressão) ou votos da maioria (classificação) para dar a resposta final

**Teoria sobre indução de regras – Algoritmos OneR e PRISM**

**Algoritmo OneR** – utiliza somente um atributo para gerar a regra.

* Testar coisas simples primeiro;
* Um atributo faz todo o trabalho;

**Algoritmo PRISM** – gera regras composta com mais de um atributo.

**Árvores x Regras**

* Compreensibilidade;
* Pouco espaço de armazenamento;
* Mais lento (regras);
* Em geral, regras não apresentam melhores resultados do que árvores de decisão.

**Outliers**

* Valores anormais;
* Valores fora do padrão (afastados da média);
* Causas:
  + Acaso;
  + Erro no preenchimento dos dados;
  + Fraudes;
* Tratamento
  + Remover o registro;
  + Não fazer nada;
  + Substituir;