aplicações que usam ML (onde eu uso isso):

* Reconhecimento facial
* kinect xbox
* jogo de futubol aprende seus movimentos e estratégias de jogo e com isso ele se comporta de uma forma para inibir suas jogadas.
* área médica para prever doenças
* sistema de recomendação (netflix, spotify, deezer, Facebook Ads)

ML é antigo (foi criado 70 anos atrás), mas começou a realmente ser aplicado e grande esforço aqui.

**Terminologia**

* IA é conceito/area mais amplo, que contem:
  + Sistemas especialistas
  + Visão computacional
  + machine learning
  + algoritmos genéticos
  + processamento de linguagem natural
  + sistema multiagente
  + mineração de dados
  + resolução de problemas por meio de busca
  + lógica fuzzy
  + robótica
  + redes neurais (é um tipo de metodo em ML)
* **IA é área que desenvolve sistemas que simula capacidade humana para resolver problemas.**
* ML são metodos para treinar algoritmos a partir de uma base de dados
* Data mining é extração conhecimentos de bases de dados, usando metodos de ML
* ML precisa de pre processamento dos dados (entregar os dados prontos para o algoritmo ser treinado e aprender). Já DM aprende de forma intuitiva (dados brutos sem pre processamento).
* Para big data precisa ser algoritmos de ML, não pode ser comuns

Pelo que foi falado na aula 6 as quatro principais tarefas de mineração de dado se encaixaria nos 2 maiores grupos/categorias da forma que está descrito abaixo

Predição (uso dados históricos para prever algo que pode acontecer no futuro)

* Regressão, o resultado (predição) é números.
* Classificação, o resultado (predição) é rótulos
  + Precisa ter uma dataset para aprender, esse dataset é uma base histórica na qual cada registro é um elemento que tem suas caracteristícas (atributos preditores) e possui sua classe/resultado (atributo meta.
  + O atributo meta é conhecido para aprendizado
  + Nos dados de entrada que o algoritmo busca prever ele baseado nas características desse elemente ele vai prever o meta.
  + Exemplo
    - Supervisor forneçe um dataset com informações sobre os clientes de uma livraria, no qual tem como elementos preditores: sexo, idade e profissão, e o atributo meta sendo comprou ou não comprou.
    - Algoritmo aprende com essa base história e consegue prever se uma pessoa vai ou não comprar o livro baseado nas suas caracteristicas.

Descritivo

* Associação, resultado é informação de relacionamento (causa-connsequencia) entre dados
* agrupamento, resultado é grupos que contém itens que possuem características semelhantes.
* detecção de anomalias, resultado é identificar que eu determinado dado está fora do padrão conhecido
* Descoberta de padrões sequência, resultado é idejtificacao de sequências
* Sumarização, resultado é informação que são comuns em um volume de dados

Eu vejo claro o relacionamento entre agrupamento e sumarização, no qual sumarização complementa o agrupamento. Meio que o agrupamento por sí só não é suficiente dependend do objetivo, pois ele agrupa elementos que tem em características comuns e o de sumarização retira as informações importantes desse grupo. Por exemplo, tem um grupo dos que investem e os que não investem. No caso para entender o que tem de características das pessoas que investem eu preciso sumarizar, com isso terei o perfil (características) das pessoas que investem e isso me possibilita tomar várias ações.

Tipo de aprendizado

* Supervisionado, é o que tem figura do supervisor (pessoa) que forneçe um conjunto de dados com caracteristícas e a resposta (o que é), o elemento que tem essas característicasn, para o algoritmo aprender
* Não supervisionado, é o que não tem figura do supervisor, aprende automaticamente identificação de padrões por meio da base de dados que é a que se espera uma resposta.
* Por reforço, é aquele que se aprende pela experiência, com as interações com o ambiente. Bastante comum em SMA (sistema multi-agentes)

**O grupo de predição é aprendizado supervizionado e grupo descritivo é não supervisionado.**