

Introdução à IA

Evolução Histórica da Computação

Prof. Maurício A Dias



Possíveis Algoritmos

Agentes inteligentes
Algoritmos Genéticos
Redes Neurais Artificiais
Métodos de Busca
Árvores de Decisão
Lógica Fuzzy
Outro Método a Escolher

Como escolher tão cedo?

IA

[Winston, 1984]:

"Inteligência artificial é o estudo das ideias que permitem aos computadores serem inteligentes"

[Schalkoff, 1990]:

"É o campo de estudo que tenta explicar e simular o comportamento inteligente em termos de processos computacionais"

[Kurzweil, 1990]

"A arte de criar máquinas que executam funções que requerem inteligência quando executadas por pessoas"

Na Prática...

[Russell & Novig, 2003]:

Sistemas que agem como humanos
Teste de Turing 🇺🇸

Sistemas que pensam como humanos
Ciência Cognitiva

Sistemas que pensam racionalmente
Lógica

Sistemas que agem racionalmente
Agentes racionais – Agem sempre
racionalmente para alcançar as suas metas

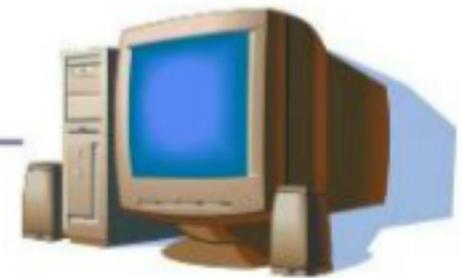


Interrogador

INTERFACE



Humano



Sistema Computacional

Na Prática...

[Russell & Novig, 2003]:

Sistemas que agem como humanos
Teste de Turing 

Sistemas que pensam como humanos
Ciência Cognitiva

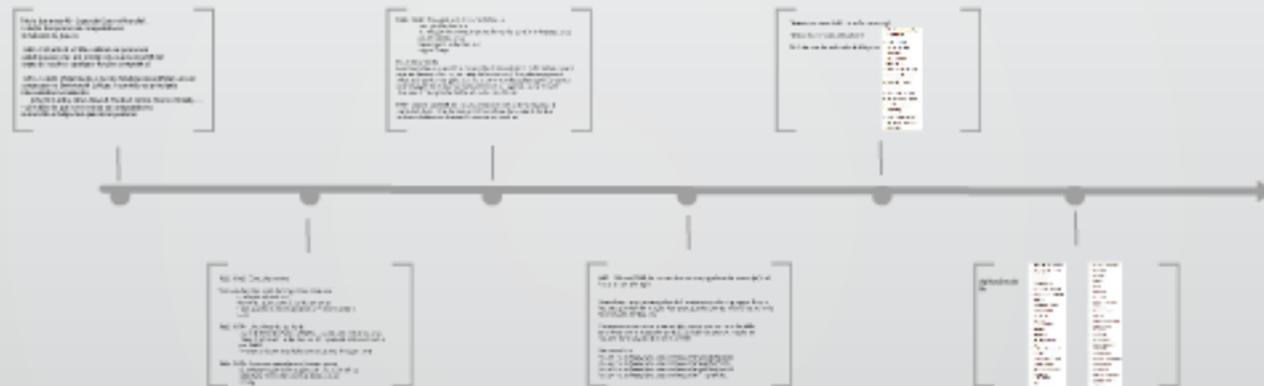
Sistemas que pensam racionalmente
Lógica

Sistemas que agem racionalmente
Agentes racionais – Agem sempre
racionalmente para alcançar as suas metas





Histórico da IA



Início dos anos 40 - Segunda Guerra Mundial.
Criação dos primeiros computadores.
Simulação de guerra.

1943 - McCulloch e Pitts realizam os primeiros estudos para criar um modelo de neurônio artificial capaz de resolver qualquer função computável.

1956 - Criado oficialmente o termo Inteligência Artificial em um congresso no Dartmouth College. Reunindo os principais interessados no assunto:

John McCarthy, Allen Newell, Herbert Simon, Marvin Minsky, ...
– previsão de que em 10 anos os computadores seriam tão inteligentes quanto as pessoas

1956 - 1966 - Época de sonhos.

Tentativa de construção de programas universais:

- tradução automática
- demonstração automática de teoremas
- resolução de problemas (General Problem Solver)
- Lisp.

1966 - 1974 - Uma dose de realidade.

Livro de Minsky e Papert (1969) critica perceptrons e paralisa investimentos em redes neurais por quase 20 anos (tema volta em 1986).

Problema da complexidade computacional dos algoritmos.

1969 - 1979 - Sistemas baseados em Conhecimento

Grande evolução da Inteligência Artificial Simbólica.

Desenvolvimento de sistemas especialistas.

Prolog.

1980 - 1988 - Inteligência Artificial na Indústria

- restrição do domínio
- introdução do conhecimento e forma de raciocínio do especialista
- Sistema especialista.
- Ressurgem redes neurais.
- Lógica Fuzzy.

Início dos anos 90:

Sistemas Especialistas com alto custo de manutenção. Erro foi não ver que o objetivo deve ser Auxiliar, ao invés de Automatizar. Grandes avanços em todas as áreas da inteligência artificial, com manifestações significativas na aprendizagem de máquina, planejamento multi-agente, raciocínio com incerteza, mineração de dados, entre outros tópicos.

1997 - Deep Blue (IBM) derrota o campeão mundial de xadrez (Garry Kasparov). Algoritmos de busca, computadores de alta velocidade e hardware desenvolvido especificamente para xadrez.

2011 – Watson (IBM) derrota os dois maiores jogadores de Jeopardy (Brad Rutter e Ken Jennings).

Baseado em técnicas avançadas de Processamento de Linguagem Natural, Recuperação de Informação, Representação de Conhecimento, Raciocínio e Aprendizado de Máquina.

Processamento paralelo massivo. 90 clusters com um total de 2880 servidores com processadores de 3.5 GHz (8 núcleos e 4 threads por núcleo). 16 Terabytes de memória RAM.

Documentário:

Parte 1: <http://www.youtube.com/watch?v=5Gpaf6NaUEw>

Parte 2: <http://www.youtube.com/watch?v=6ay17a7mEIk>

Parte 3: <http://www.youtube.com/watch?v=gphA9u5nm5U>

Parte 4: <http://www.youtube.com/watch?v=ilrKOovFpVc>

Novos caminhos da IA - <http://aitopics.org/>

Nossa disciplina está focada em:

IA Clássica e Aprendizado de Máquina

> **Machine Learning
Overview**

> Data Mining

> Decision Tree
Learning

> Evolutionary
Algorithms

> Inductive Learning

> Learning Theory

> Memory-Based
Learning

> Neural Networks

> Pattern Recognition

> Reinforcement
Learning

> Scientific Discovery

> Statistical Learning
Methods

Aplicações da IA

- > Assistive Technologies
- > Astronomy & Space Exploration
- > Automatic Programming
- > Automotive Industry
- > Autonomous Vehicles
- > Aviation
- > Banking & Finance
- > Bioinformatics
- > Biometrics
- > Business & Manufacturing
- > Chatbots
- > Chemistry
- > Decision Support Systems
- > Earth & Atmospheric Science
- > Engineering Design
- > Fraud Detection
- > Hazards & Disasters
- > Knowledge Management
- > Law
- > Law Enforcement &

- > Machine Storytelling
- > Marketing
- > Medicine
- > Military
- > Music
- > Networks
- > Oil & Gas
- > Politics & Foreign Relations
- > Recommender Systems
- > Robots in the Home
- > Robots in the Workplace
- > Science & Mathematics
- > Smart Houses & Appliances
- > Social Science
- > Software Engineering
- > Spam Filtering
- > Surveillance
- > Telecommunications
- > Transportation & Shipping

Dado, Informação, Conhecimento

- ❑ Dado: é a estrutura fundamental sobre a qual um sistema de informação é construído
- ❑ Informação: a transformação de dados em informação é freqüentemente realizada através da apresentação dos dados em uma forma compreensível ao usuário
- ❑ Conhecimento:
 - Fornece a capacidade de resolver problemas, inovar e aprender baseado em experiências prévias
 - Uma combinação de instintos, idéias, regras e procedimentos que guiam as ações e decisões



REPRODUZIR A
INTELIGÊNCIA HUMANA



- (2) Definição de Inteligência: Faculdade de conhecer, de aprender, de conceber, de compreender: a inteligência distingue o homem do animal [Larousse 99];
- (3) Definição de Inteligência Artificial: Conjunto de teorias e de técnicas empregadas com a finalidade de desenvolver máquinas capazes de simular a inteligência humana [Larousse99];
- (4) Definição de Inteligência Artificial: A Inteligência Artificial é uma área de estudos da computação que se interessa pelo estudo e criação de sistemas que possam exibir um comportamento inteligente e realizar tarefas complexas com um nível de competência que é equivalente ou superior ao de um especialista humano [Nikolopoulos 97].

- As definições são usualmente recursivas, ou seja, um ser inteligente é aquele que possui inteligência, e a inteligência é a característica dos seres inteligentes;
- As definições possuem contradições, onde aparece que a inteligência distingue o homem do animal, pois só o homem é capaz de aprender, conhecer, compreender. Isto não é muito preciso/correto, pois os animais também podem aprender (a reconhecer o dono), conhecer (o caminho de volta para casa) e compreender (um comando de seu mestre). E podemos mesmo dizer que estas tarefas realizadas por animais são tipicamente tarefas consideradas na atualidade como “tarefas inteligentes” que os computadores ainda não realizam com um desempenho satisfatório.
- Uma pessoa com limitações (analfabeta, deficiente mental e/ou físico, com QI reduzido) poderá ser qualificada, pelos critérios acima, como inteligente? E no entanto ela é um ser inteligente com capacidades superiores as de um computador, ou você acha que, por exemplo, um computador conseguiria ir até uma farmácia comprar um remédio? (isto inclui tarefas de: planificação de trajetória, reconhecimento do caminho, tratamento de imprevistos, etc).

- Associação de idéias e conceitos;
- Concluir coisas;
- Capacidade de aprendizado;
- Acúmulo de conhecimentos;
- Raciocínio: lógico, abstrato, dedução, analogia, indução, inferência, síntese, análise, ...;
- Uso de experiências e conhecimentos passados;
- Uso prático de conhecimentos e experiências;
- Tomada de decisões;
- Criar coisas novas (criatividade);
- Saber o que eu sei (capacidade de explicar algo);
- Saber que sei / Saber que não sei;
- Interação;
- Comunicação.

Vamos tentar definir aqui brevemente o que entendemos por aprendizado. Aprendizado é a capacidade de se adaptar, de modificar e melhorar seu comportamento e suas respostas, sendo portanto uma das propriedades mais importantes dos seres ditos inteligentes, sejam eles humanos ou não.

- **Adaptação** e mudança de comportamento de forma a evoluir (melhorar segundo algum critério). Um sistema, biológico ou artificial, que não seja capaz de evoluir ou de mudar seu comportamento diante de novas situações que lhe são propostas é um sistema sem inteligência;
- **Correção** dos erros cometidos no passado, de modo a não repeti-los no futuro. Este item está diretamente relacionado ao anterior, o sistema deve modificar seu comportamento caso o comportamento atual não satisfaça a algum tipo de exigência (onde a sobrevivência deve ser um dos quesitos mais importantes a serem considerados nos seres vivos);
- **Otimização**: melhoria da performance do sistema como um todo. O aprendizado pode implicar em uma mudança do comportamento que busque: a economia de energia gasta para realizar uma tarefa, a redução do tempo gasto numa tarefa, etc. Quando falamos em otimização, devemos lembrar que podemos ter quesitos contraditórios e opostos, onde teremos que maximizar ou minimizar custos de acordo com algum tipo de critério;
- **Interação** com o meio, pois é através deste contato com o mundo que nos cerca que podemos trocar experiências e/ou realizar experiência, de forma a adquirir novos conhecimentos;
- **Representação** do conhecimento adquirido. O sistema deve ser capaz de armazenar uma massa muito grande de conhecimentos, e isto requer uma forma de representar estes conhecimentos que permita ao sistema explorá-los de maneira conveniente. Como nossos recursos são limitados, devemos ter uma maneira de guardar conhecimentos e regras gerais, pois guardar tudo seria muito difícil (exige muita memória, dificulta o uso pela lentidão da consulta aos conhecimentos).

- 
- **Adaptação** e mudança de comportamento de forma a evoluir (melhorar segundo algum critério). Um sistema, biológico ou artificial, que não seja capaz de evoluir ou de mudar seu comportamento diante de novas situações que lhe são propostas é um sistema sem inteligência;
 - **Correção** dos erros cometidos no passado, de modo a não repeti-los no futuro. Este item está diretamente relacionado ao anterior, o sistema deve modificar seu comportamento caso o comportamento atual não satisfaça a algum tipo de exigência (onde a sobrevivência deve ser um dos quesitos mais importantes a serem considerados nos seres vivos);
 - **Otimização**: melhoria da performance do sistema como um todo. O aprendizado pode implicar em uma mudança do comportamento que busque: a economia de energia gasta para realizar uma tarefa, a redução do tempo gasto numa tarefa, etc. Quando falamos em otimização, devemos lembrar que podemos ter quesitos contraditórios e opostos, onde teremos que maximizar ou minimizar custos de acordo com algum tipo de critério;
 - **Interação** com o meio, pois é através deste contato com o mundo que nos cerca que podemos trocar experiências e/ou realizar experiência, de forma a adquirir novos conhecimentos;

- **Correção** dos erros cometidos no passado, de modo a não repeti-los no futuro. Este item está diretamente relacionado ao anterior, o sistema deve modificar seu comportamento caso o comportamento atual não satisfaça a algum tipo de exigência (onde a sobrevivência deve ser um dos quesitos mais importantes a serem considerados nos seres vivos);
- **Otimização**: melhoria da performance do sistema como um todo. O aprendizado pode implicar em uma mudança do comportamento que busque: a economia de energia gasta para realizar uma tarefa, a redução do tempo gasto numa tarefa, etc. Quando falamos em otimização, devemos lembrar que podemos ter quesitos contraditórios e opostos, onde teremos que maximizar ou minimizar custos de acordo com algum tipo de critério;
- **Interação** com o meio, pois é através deste contato com o mundo que nos cerca que podemos trocar experiências e/ou realizar experiência, de forma a adquirir novos conhecimentos;
- **Representação** do conhecimento adquirido. O sistema deve ser capaz de armazenar uma massa muito grande de conhecimentos, e isto requer uma forma de representar estes conhecimentos que permita ao sistema explorá-los de maneira conveniente. Como nossos recursos são limitados, devemos ter uma maneira de guardar conhecimentos e regras gerais, pois guardar tudo seria muito difícil (exige muita memória, dificulta o uso pela lentidão da consulta aos conhecimentos).

Sistemas Especialistas - Esquema Global

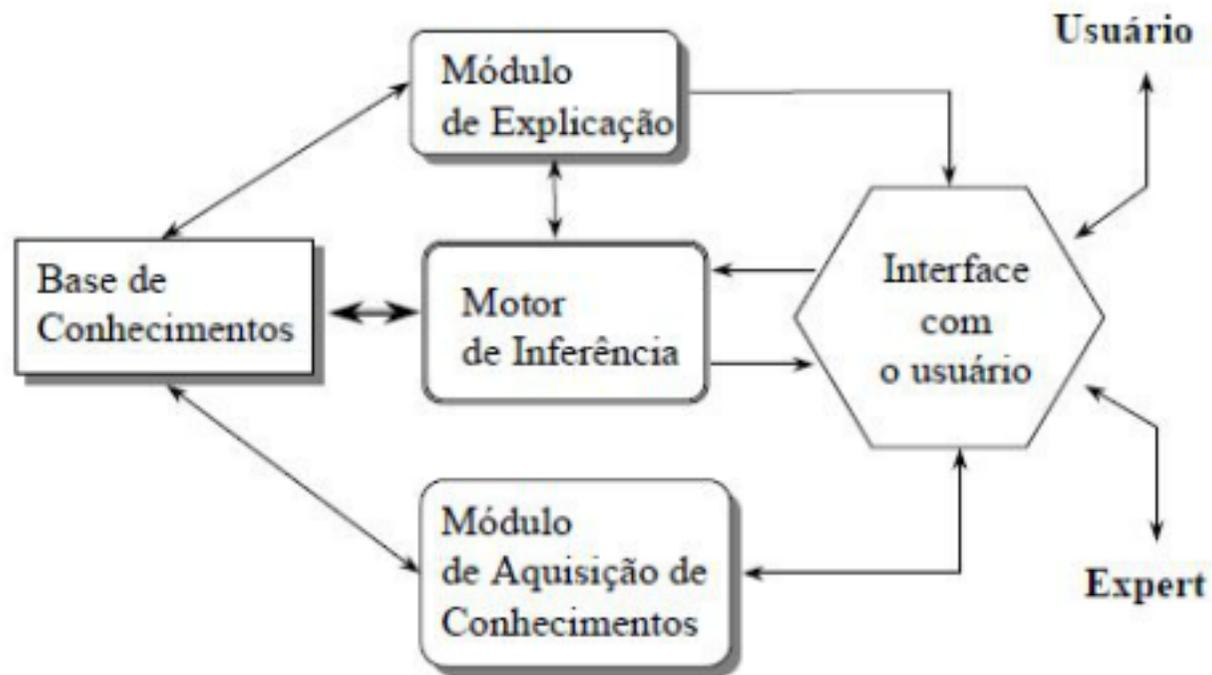


Figura 1.2 – Arquitetura dos Sistemas Especialistas

Base de
Conhecimentos
(regras + fatos)

Conversão para um formato
de representação interno

Aquisição
Automática

Explicitação

Conhecimentos sobre uma área de aplicação:

- *Conhecimentos do especialista*
- *Teorias sobre o domínio de aplicação*
- *Resultados de experiências praticas
(casos conhecidos)*



Inteligência Artificial

Sistemas Inteligentes ←

Aprendizado de Máquinas ←

Representação de Conhecimentos ←

Sistemas Especialistas
KBS, robótica,
visão artificial, ...

CBR, ILP, árvores de
decisão/indução, redes neurais
algoritmos genéticos, ...

Métodos Simbólicos
Métodos Conexionistas (RNA)

- Aprendizado por analogia/por instâncias. Exemplo: Sistemas baseados em casos - CBR (*Case Based Reasoning*) [Mitchell 97, Kolodner 93];
- Aprendizado por Indução. Exemplos: Árvores de Decisão - ID3, C4.5, CN2 (IDT - *Induction of Decision Trees*) [Quinlan 92], e ILP - *Inductive Logic Programming (Prolog)* [Nilsson 98].
- Aprendizado por evolução/seleção. Exemplo: Algoritmos Genéticos - GA e GP (*Genetic Algorithms / Genetic Programming*) [Goldberg 89, Mitchell 97];
- Outros tipos de aprendizado: por reforço (*reinforcement learning*), não supervisionado, bayesiano e por explicações (*explanation based*) [Mitchell 97, Nilsson 98].

□ **Conceitos de aprendizado de máquina**

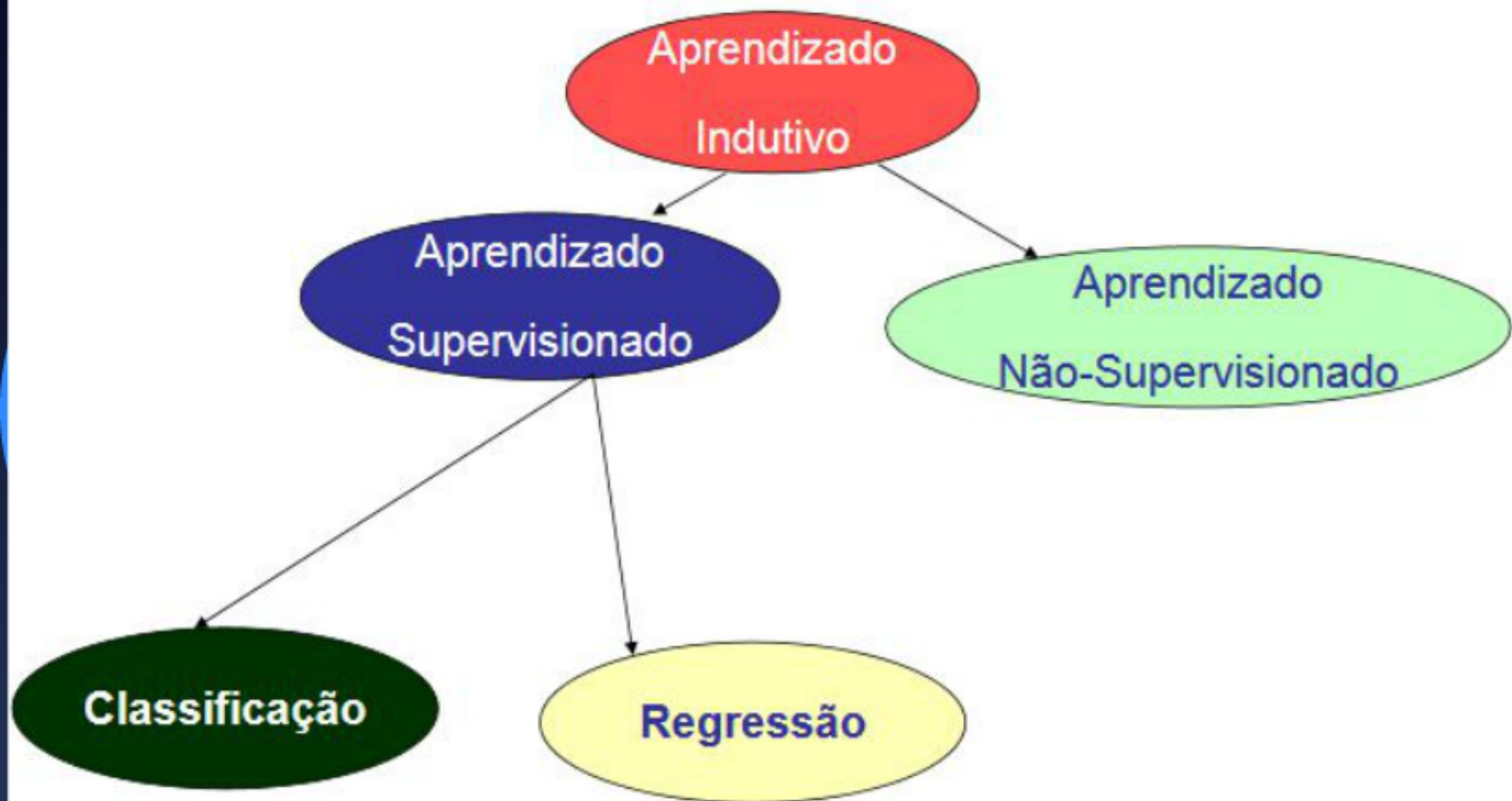
Sistemas de IA

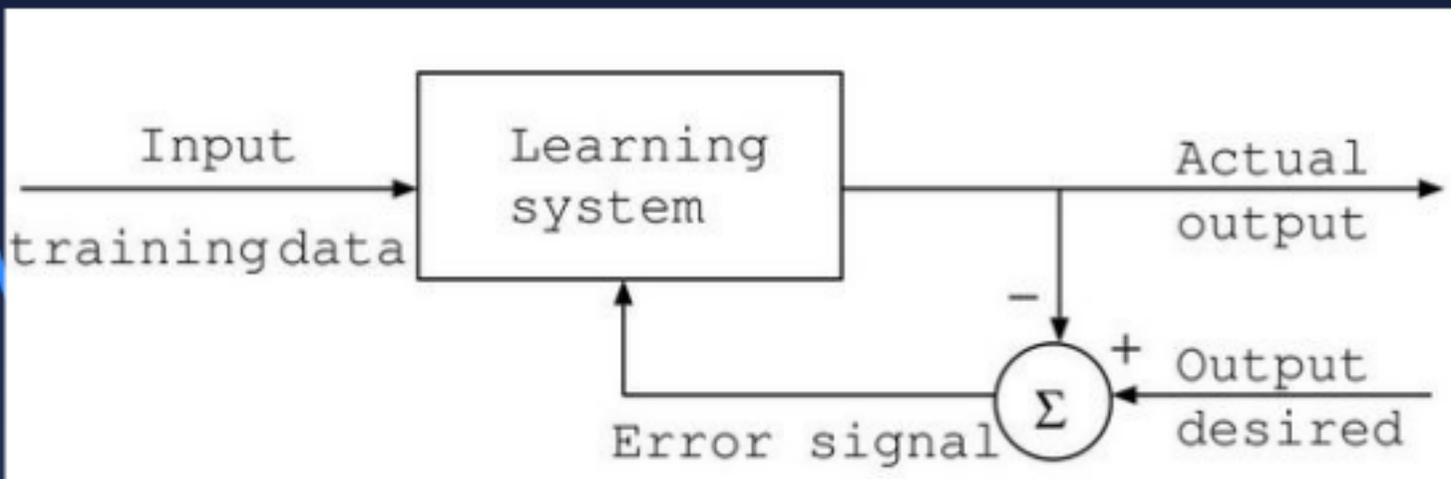
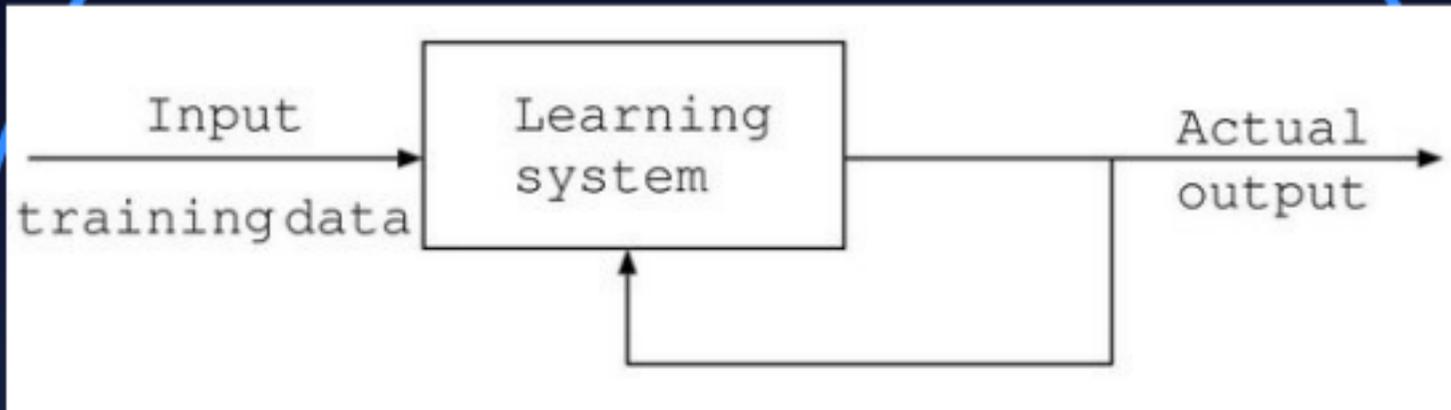
- Armazenar conhecimento ⇒ **Representação**
- Aplicar conhecimento para resolver problemas
 - **Raciocínio (mecanismo de inferência)**
- Adquirir novos conhecimentos ⇒ **Aprendizagem**

□ O que é **Aprendizado de Máquina**?

- É uma área de IA cujo objetivo é o **desenvolvimento de técnicas computacionais** sobre o **aprendizado** bem como a construção de sistemas capazes de **adquirir conhecimento de forma automática**.
- Um **sistema de aprendizado** é um programa de computador que toma decisões baseado em experiências acumuladas por meio de solução bem sucedida de problemas anteriores.
- É uma **ferramenta poderosa para aquisição automática** de conhecimento, entretanto, **não existe um único algoritmo** que apresente **melhor desempenho** para todos problemas.

- ❑ **Indução:** é a forma de inferência lógica que permite obter conclusões a partir de um conjunto de exemplos.
- ❑ Na indução, um conceito é aprendido efetuando-se **inferência indutiva** sobre os **exemplos apresentados** (*cautela na escolha de exemplos*).
- ❑ **Dedução:** Humanos usam raciocínio dedutivo para deduzir nova informação a partir de informação relacionada logicamente.





▣ Paradigmas do aprendizado:

- **Simbólico:** Buscam aprender construindo representações simbólicas (expressão lógica, **árvores de decisão regras**).
- **Estatístico:** Buscam métodos estatísticos (**Aprendizado bayesiano**)
- **Baseado em Exemplos:** Sistemas *lazy* (RBC, **Nearest Neighbors**).
- **Conexionista:** Modelos inspirados no modelo biológico do sistema nervoso (**Redes Neurais**).
- **Evolutivo:** Teoria de Darwin (**Algoritmos Genéticos**).

□ Algumas Definições em AM

- **Bias:** qualquer preferência de uma hipótese sobre a outra.
- **Modo de aprendizado:**
 - todo conjunto de treinamento presente no aprendizado (**não incremental**).
 - quando novos exemplos de treinamento são adicionados (**incremental**).

Erro ($err(h)$)

Medida de desempenho de um Classificador.

Considerando $\|E\| = \begin{cases} 1 & \text{se a expressão for verdadeira} \\ 0, & \text{caso contrário} \end{cases}$

$$err(h) = \frac{1}{n} \sum_{i=1}^n \| y_i \neq h(x_i) \|$$

Acurácia ($acc(h)$)

Complemento da Taxa de Erro, representa a Precisão do Classificador.

$$acc(h) = 1 - err(h)$$

Distribuição de Classes ($distr(C_j)$)

Para cada Classe C_j , sua distribuição $distr(C_j)$ é calculada como sendo o número de exemplos em T que possuem classe C_j dividido pelo número total de exemplos (n), ou seja, a proporção de exemplos em cada classe

$$distr(C_j) = \frac{1}{n} \sum_{i=1}^n \mathbb{1} \{ y_i = C_j \}$$

Exemplo: Conjunto com 100 Exemplos: 60 Classe C1
15 Classe C2
25 Classe C3

$$distr(C1, C2, C3) = (60\%, 15\%, 25\%)$$

Neste exemplo, **Classe Majoritária** (ou **Prevalente**) é C1.
Classe **Minoritária** é C2.

Erro Majoritário ($maj-err(T)$)

Limite Máximo abaixo do qual o erro de um Classificador deve ficar

$$maj - err(T) = 1 - \max_{i=1, \dots, k} distr(C_i)$$

No Exemplo anterior: $maj-err(T) = 1 - 0,60 = 0,40$

Erro Majoritário **INDEPENDENTE** do algoritmo de aprendizado.

Prevalência de Classe

Problema com **desbalanceamento** de classes em conjunto de exemplos.

Exemplo: $\text{distr}(C1, C2) = (99,75\%, 0,25\%)$

Neste exemplo, Classe **Majoritária** (ou **Prevalente**) é C1
Classe **Minoritária** é C2

Classificador que classifique novos exemplos como C1 teria uma precisão de 99,75%.

Se a Classe C2 fosse, por exemplo, ocorrência de Geada ...

“Overfitting”

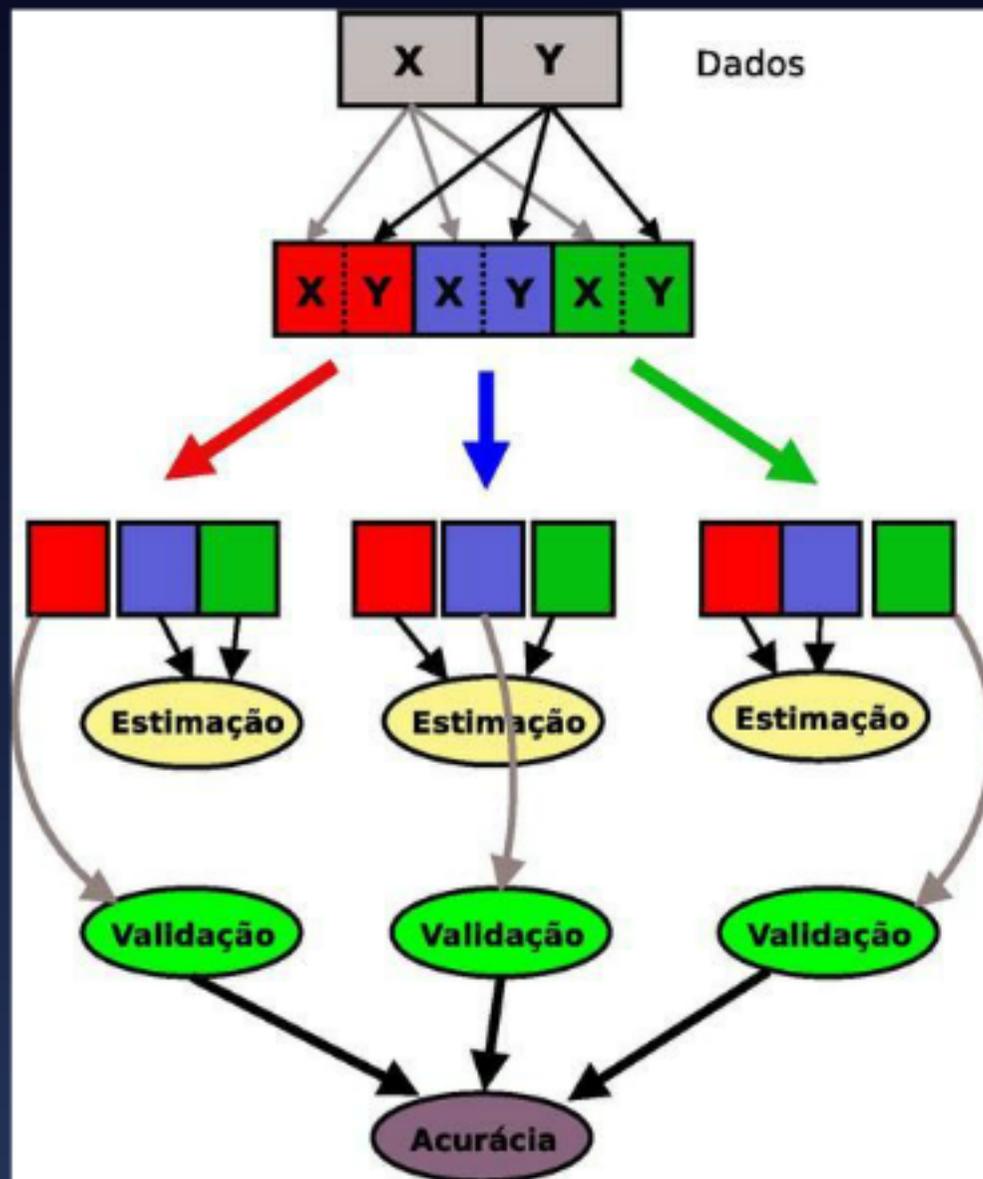
- É possível que o Classificador faça uma indução muito específica para o conjunto de treinamento utilizado (“Overfitting”).
- Como este é apenas uma amostra de dados, é possível que a indução tenha bom desempenho no conjunto de treinamento, mas um desempenho ruim em exemplos diferentes daqueles pertencentes ao conjunto de treinamento.
- **Cálculo do Erro** em um conjunto de teste independente evidencia a situação de “Overfitting”.
- **Under** e **overfitting**: ajusta-se em muito pouco ou em excesso ao conjunto de treinamento.

Underfitting

- ❑ A hipótese induzida apresenta um desempenho ruim tanto no conjunto de treinamento como de teste
 - poucos exemplos representativos foram dados ao sistema de aprendizado (e.g. algoritmos de árvores de decisão ou de indução de regras)
 - o usuário pré-definiu um tamanho muito pequeno para o classificador (e.g. insuficientes neurônios em uma rede neural ou um alto valor de poda para árvores de decisão)

- Dado um conjunto de registros (dataset):
 - Cada registro contém um conjunto de **atributos**, em que um dos atributos é o **atributo-meta** (**variável resposta**).
 - O conjunto de dados é dividido em dois subconjuntos: **conjunto de treinamento** para construir o modelo e **conjunto de teste** para validar o modelo.
- **Passo 1**: encontrar um **modelo** para o atributo-meta (ou atributo-classe) como uma função dos valores dos outros atributos.
- **Passo 2**: registros não conhecidos devem ser associados à classe com a maior precisão possível.

K-FOLD



Observação Importante

- **Qualidade** do Exemplo de Treinamento → **Qualidade das Regras.**
- **Não** é possível **descobrir** algo que **não** esteja nos exemplos.
- **Seleção dos exemplos** e das **características** é fundamental neste processo.
 - Daí a **importância** do **conhecimento** e da **experiência** do **Especialista.**

Características de um bom classificador

- **Precisão**
- **Velocidade**
 - Tempo para construir o modelo.
 - Tempo para usar o modelo.
- **Robustez**
 - Capacidade de lidar com ruídos e valores faltantes (**missing**).
- **Escalabilidade**
 - Eficiência em banco de dados residentes em disco.
- **Interpretabilidade**
 - Clareza fornecida pelo modelo.
- **Relevância na seleção de regras**
 - Tamanho da árvore de decisão.
 - Regras de classificação compactas.

Ruído

- ❑ Exemplos imperfeitos que podem ser derivados do processo de aquisição, transformação ou rotulação das classes
- ❑ Ex: exemplos com os mesmos atributos mas com classes diferentes

	X ₁	X ₂	X ₃	X ₄	Y
overcast	19	65	yes	dont_go	
rain	19	70	yes	dont_go	
rain	23	80	yes	dont_go	
sunny	23	95	no	dont_go	
sunny	28	91	yes	dont_go	
sunny	30	85	no	dont_go	
overcast	19	65	yes	go	
rain	21	80	no	go	
rain	22	95	no	go	
sunny	22	70	no	go	
overcast	23	90	yes	go	
rain	25	81	no	go	
sunny	25	72	yes	go	
overcast	26	75	no	go	
overcast	29	78	no	go	

❑ Principais fatores de erro:

- Qualidade (representatividade) da informação dos atributos
- Adaptação do algoritmo de aprendizado aos exemplos
- Distribuição dos exemplos futuros
- Quantidade de exemplos