

Aula 5
Prof. Daniel Cavalcanti Jeronymo

Inteligência Computacional

CP78D

Introdução à aprendizagem de máquina. Árvores de Decisão. Sistemas Especialistas e motores de inferência.

Universidade Tecnológica Federal do Paraná (UTFPR)
Engenharia Eletrônica – 8º Período

- Introdução à aprendizagem de máquina
- Árvores de Decisão
- Sistemas Especialistas
- Motores de Inferência

Introdução à aprendizagem de máquina

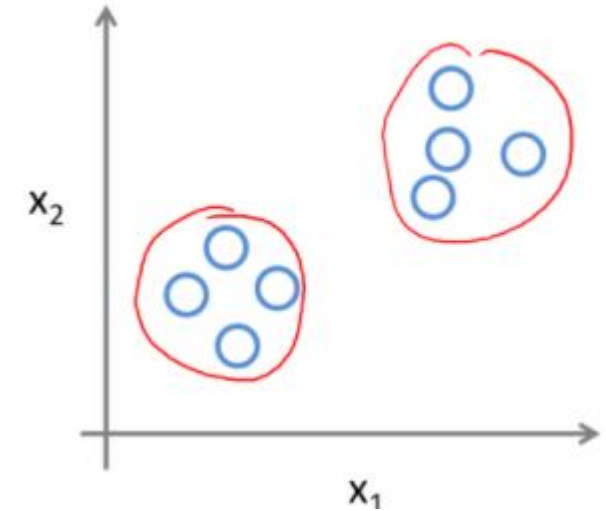
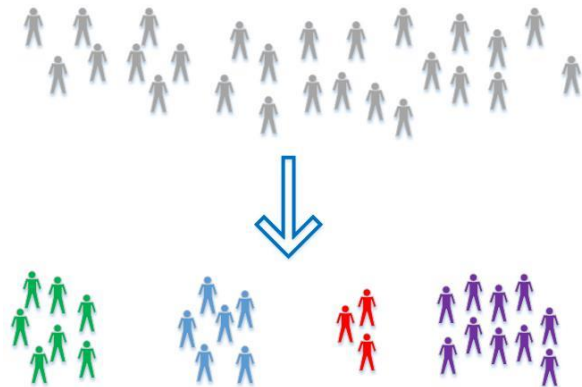
- Tipos de aprendizagem de máquina

Não-supervisionado

Conjunto de dados contém apenas (x)

Conjunto de características X_1, X_2, \dots, X_p

Objetivo: agrupar itens similares (estrutura)



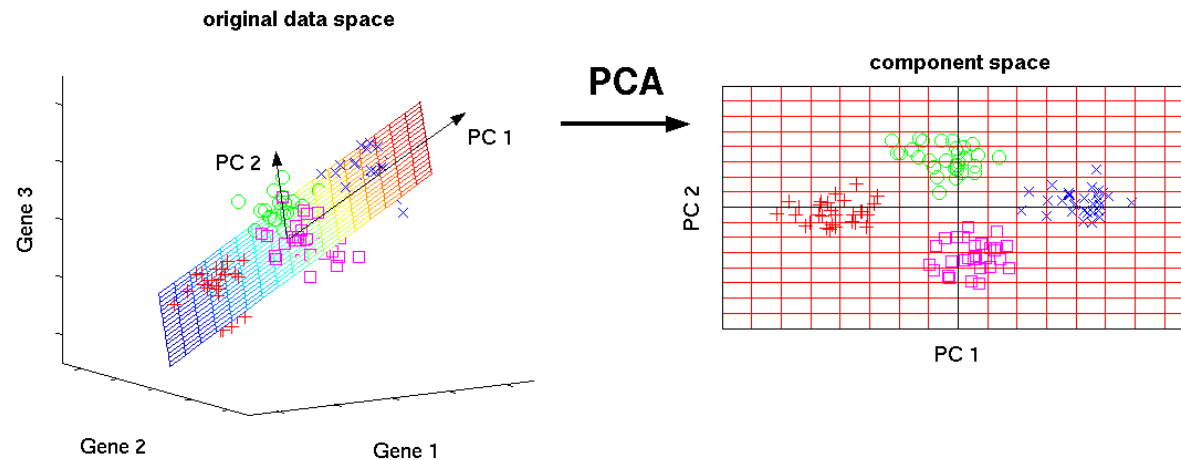
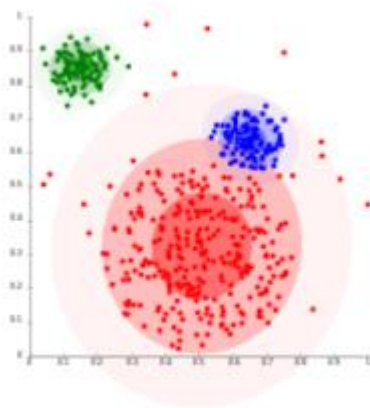
Introdução à aprendizagem de máquina

- Tipos de aprendizagem de máquina

Não-supervisionado

Clustering (e.g., k-means)

Redução de Dimensionalidade (e.g., PCA)



Introdução à aprendizagem de máquina

- Tipos de aprendizagem de máquina

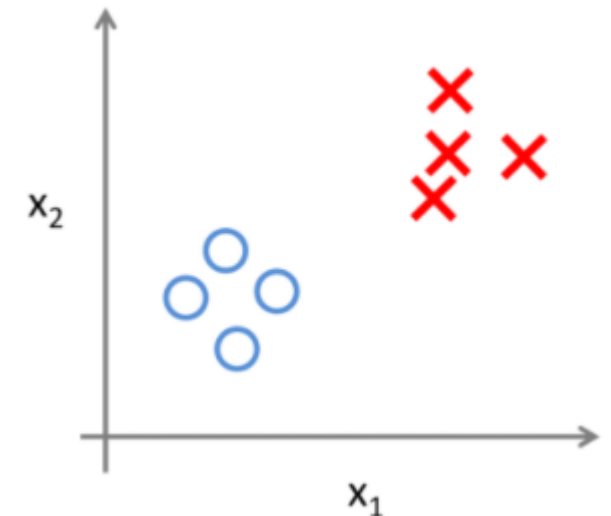
Supervisionado

Conjunto de dados contém pares de entrada-saída (x,y)

Conjunto de características X_1, X_2, \dots, X_p

Dados classificados por classes Y

Objetivo: realizar previsões de novas características



Introdução à aprendizagem de máquina


- Tipos de aprendizagem de máquina



Semisupervisionado

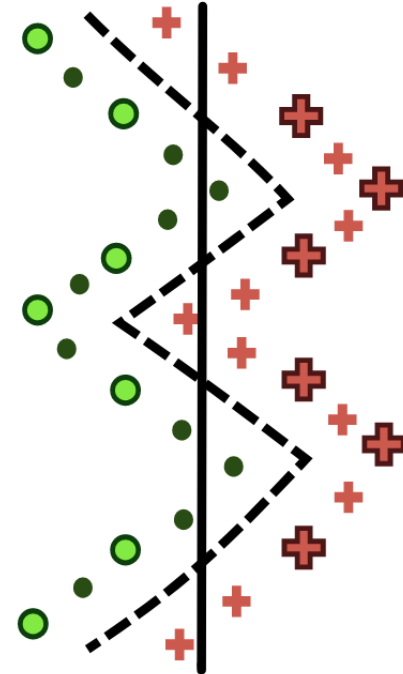
Conjunto de dados contém poucos pares (x,y)

Conjunto de dados contém muitos (x)

— Fronteira de decisão dos dados classificados
 - - - Fronteira de decisão ideal

 Classificado
 Não-classificado

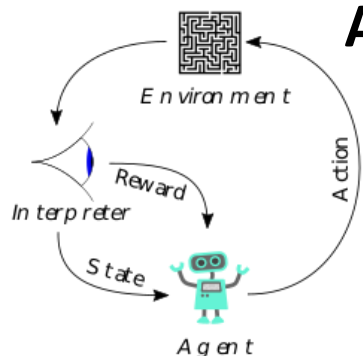
Class 2
 Classificado
 Não-classificado



Aprendizado por Reforço

Conjunto de dados contém poucos pares (x,y)

Conjunto de dados contém muitos (x)



Árvores de Decisão

- Aprendizado de máquina supervisionado

Entrada

Conjunto de atributos

Saída

Decisão ou previsão do valor de saída

Conjunto de decisões em forma de testes

Cada decisão gera uma ramificação na árvore



Árvores de Decisão

- Exemplo: necessário esperar por uma mesa num restaurante

Definir atributos que descrevem o domínio

Objetivo: será necessário esperar por uma mesa?

Árvores de Decisão

- Exemplo: necessário esperar por uma mesa num restaurante

Alternativa: Verdadeiro se existe um restaurante alternativo adequado nas proximidades.

Bar: Verdadeiro se o restaurante tem uma área de bar confortável para ficar esperando.

Sex/Sab: Verdadeiro se o dia da semana for sexta ou sábado.

Faminto: Verdadeiro se estamos com fome.

Pessoas: Quantas pessoas estão no restaurante (os valores são Nenhuma, Algumas e Cheio).

Preço: Preço do restaurante de (\$, \$ \$, \$\$\$).

Chuva: Verdadeiro se está chovendo lá fora.

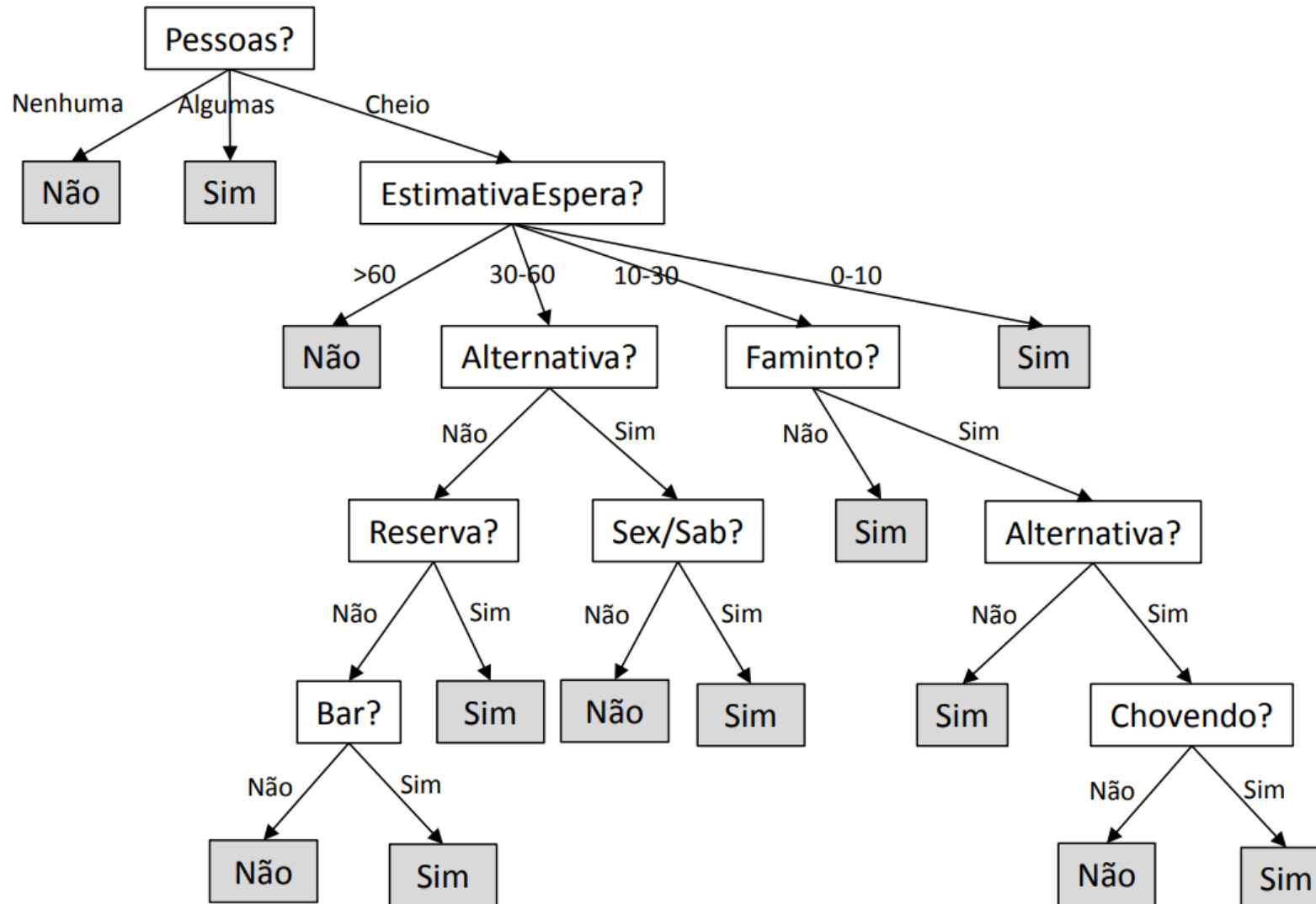
Reserva: Verdadeiro se nós fizemos uma reserva.

Tipo: Tipo de restaurante (Francês, Italiano, Tailandês, Hambúrguer).

EstimativaEspera: Tempo de espera estimado (00-10, 10-30, 30-60, > 60 minutos).

Árvores de Decisão

- Exemplo: necessário esperar por uma mesa num restaurante



Árvores de Decisão

- Exemplo: necessário esperar por uma mesa num restaurante
- A árvore de decisão pode ser obtida arbitrariamente ou criada a partir de exemplos

	Atributos										Obj.
Exemplo	Alt.	Bar	S/S	Fam.	Pes.	Pre.	Chov.	Res.	Tipo	Est.	Esp.
X_1	Sim	Não	Não	Sim	Algumas	\$\$\$	Não	Sim	Fran.	0-10	Sim
X_2	Sim	Não	Não	Sim	Cheio	\$	Não	Não	Tai.	30-60	Não
X_3	Não	Sim	Não	Não	Algumas	\$	Não	Não	Ham.	0-10	Sim
X_4	Sim	Não	Sim	Sim	Cheio	\$	Sim	Não	Tai.	10-30	Sim
X_5	Sim	Não	Sim	Não	Cheio	\$\$\$	Não	Sim	Fran.	>60	Não
X_6	Não	Sim	Não	Sim	Algumas	\$\$	Sim	Sim	Ital.	0-10	Sim
X_7	Não	Sim	Não	Não	Nenhuma	\$	Sim	Não	Ham.	0-10	Não
X_8	Não	Não	Não	Sim	Algumas	\$\$	Sim	Sim	Tai.	0-10	Sim
X_9	Não	Sim	Sim	Não	Cheio	\$	Sim	Não	Ham.	>60	Não
X_{10}	Sim	Sim	Sim	Sim	Cheio	\$\$\$	Não	Sim	Ital.	10-30	Não
X_{11}	Não	Não	Não	Não	Nenhuma	\$	Não	Não	Tai.	0-10	Não
X_{12}	Sim	Sim	Sim	Sim	Cheio	\$	Não	Não	Ham.	30-60	Sim

Árvores de Decisão

- Exemplo: necessário esperar por uma mesa num restaurante

Aprendizado de uma árvore de decisão é NP-completo

Difícil!!!

Múltiplos algoritmos

CART (Classification And Regression Tree)

ID3 (Iterative Dichotomiser 3)

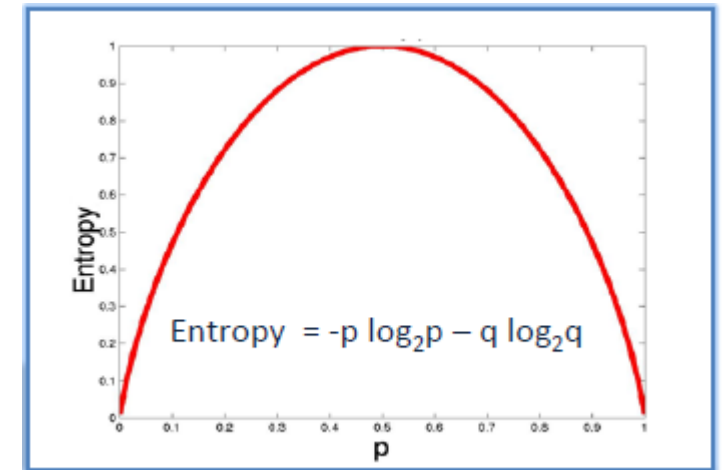
C4.5

CHAID (Chi-Squared Automatic Interaction Detector)

MARS

Conditional Inference Trees

Etc...



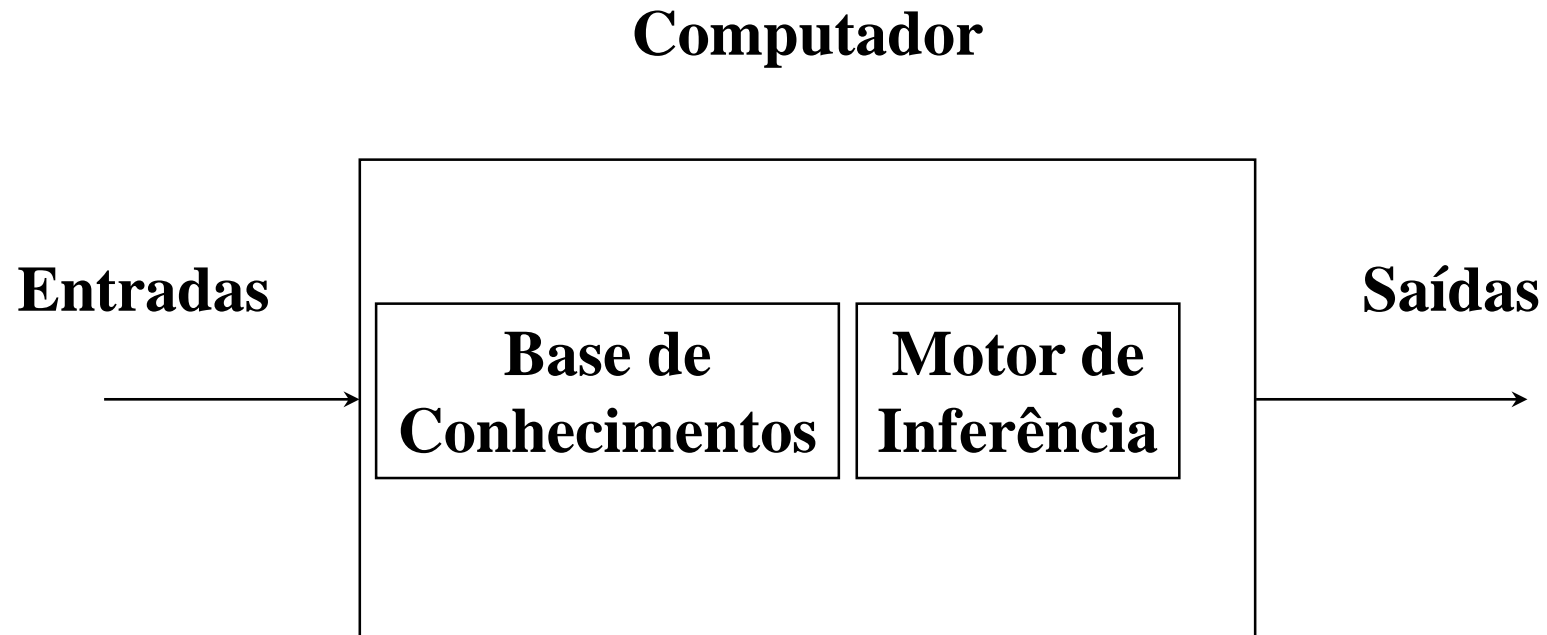
Entropia nula: conjunto homogêneo

Entropia máxima: conjunto diverso

Outras métricas: ganho de informação (GINI)

Sistemas Especialistas

- Sistemas com conhecimento e raciocínio básico



Sistemas Especialistas

- Sistemas com conhecimento e raciocínio básico

Base de Conhecimentos (Fatos)

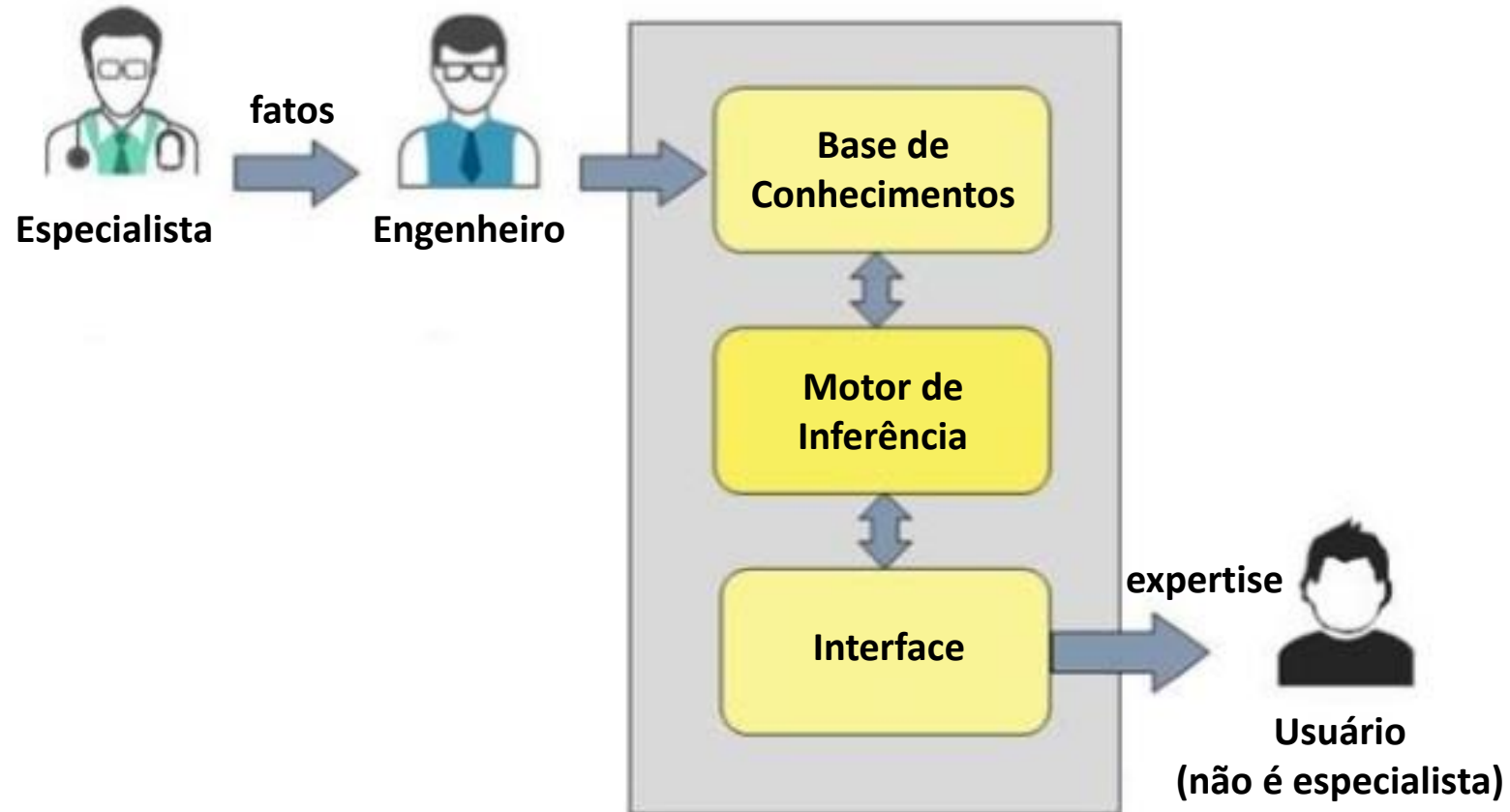
Obtida à partir de livros, pessoas especialistas num assunto, etc...

Motor de Inferência

Descobre relações entre os conhecimentos

Sistemas Especialistas

- Sistemas com conhecimento e raciocínio básico

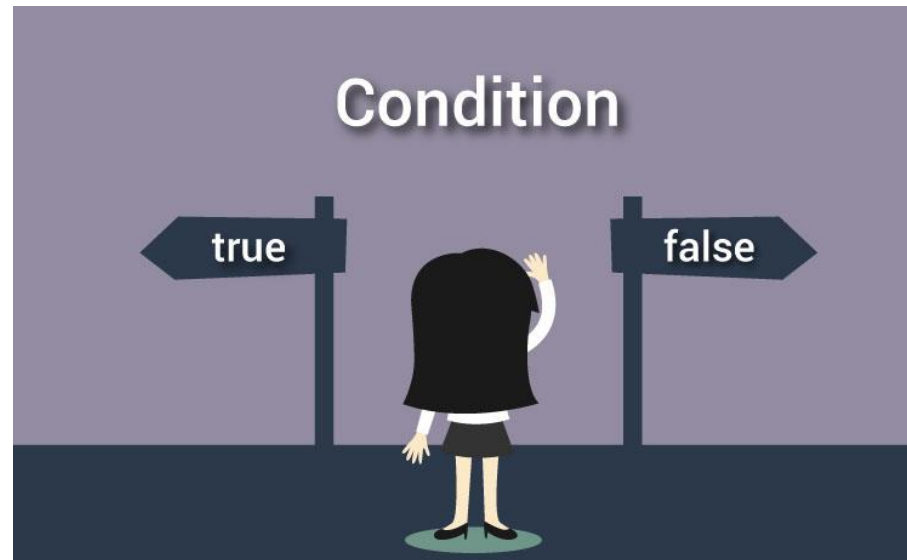


Sistemas Especialistas

- Representação do conhecimento

Regras **SE-ENTÃO**

SE você está com fome **ENTÃO** coma



Sistemas Especialistas

- Motor de Inferência

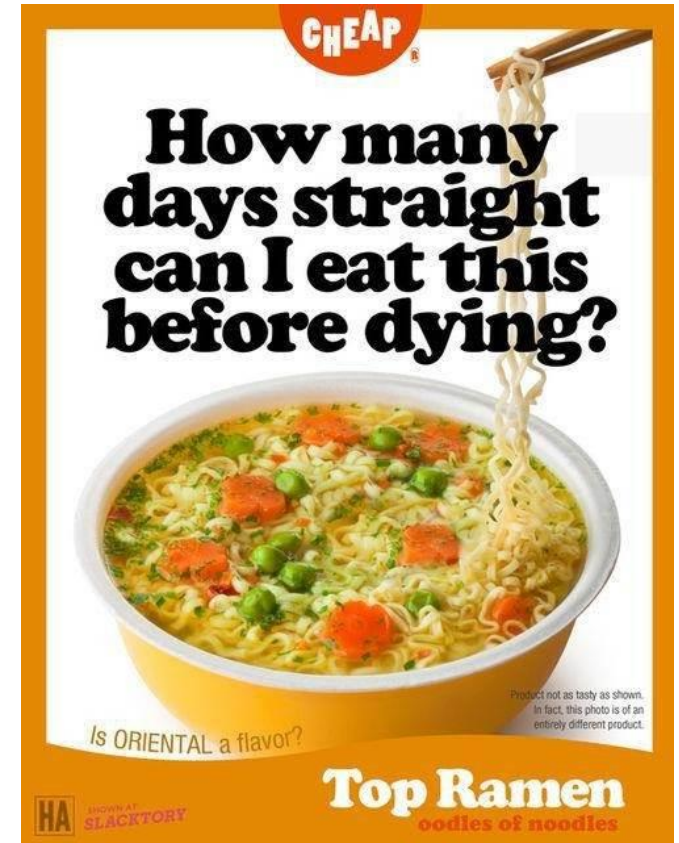
Permite encadeamento de regras

SE você está com fome **ENTÃO** coma

SE não tem comida **ENTÃO** compre

SE mercado & dinheiro < 5 reais **ENTÃO** miojo

SE mercado & dinheiro >= 5 reais **ENTÃO** macarrão



Sistemas Especialistas

- Motor de Inferência

Encadeamento à frente (*Forward Chaining*)

Fatos iniciais

Regras são aplicadas para gerar todas as conclusões

Corresponde à *modus ponens*

Orientado à dados (*data driven*)

Encadeamento para trás (*Backward Chaining*)

Conclusões iniciais

Procura-se regras que gerem as conclusões

Corresponde à *modus tollens*

Orientado à objetivos (*goal directed*)

Sistemas Especialistas

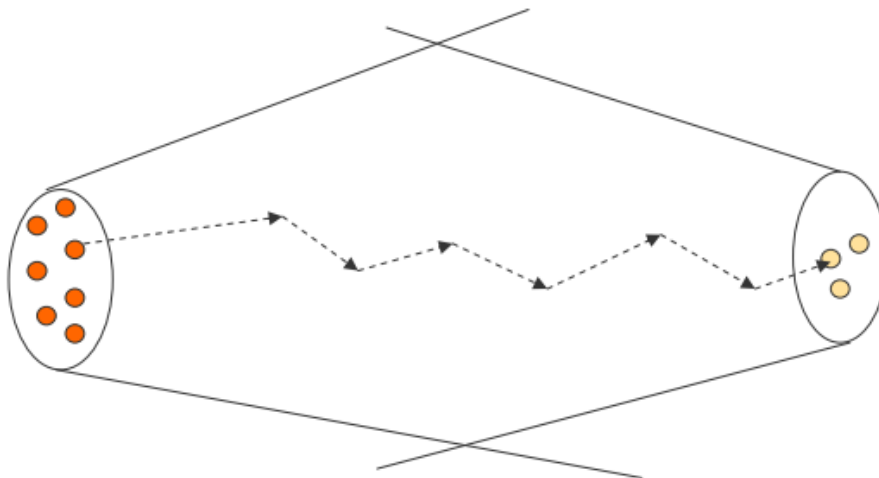
- Motor de Inferência

Encadeamento à frente (*Forward Chaining*)

Fatos iniciais levam à conclusões

Encadeamento para trás (*Backward Chaining*)

Conclusões iniciais levam à fatos



Sistemas Especialistas

- Motor de Inferência



SE X anda em teia **ENTÃO** X é uma aranha

SE X voa e X não anda em teia **ENTÃO** X é um inseto

SE X é um inseto **ENTÃO** X tem asas

SE X é uma aranha **ENTÃO** X tem 8 pernas



Aquecimento

- Base de dados Titanic

<https://www.kaggle.com/c/titanic/data>

titanic.py e **titanic_train.csv**

Data Dictionary

Variable	Definition
survival	Survival
pclass	Ticket class
sex	Sex
Age	Age in years
sibsp	# of siblings / spouses aboard the Titanic
parch	# of parents / children aboard the Titanic
ticket	Ticket number
fare	Passenger fare
cabin	Cabin number
embarked	Port of Embarkation

Key

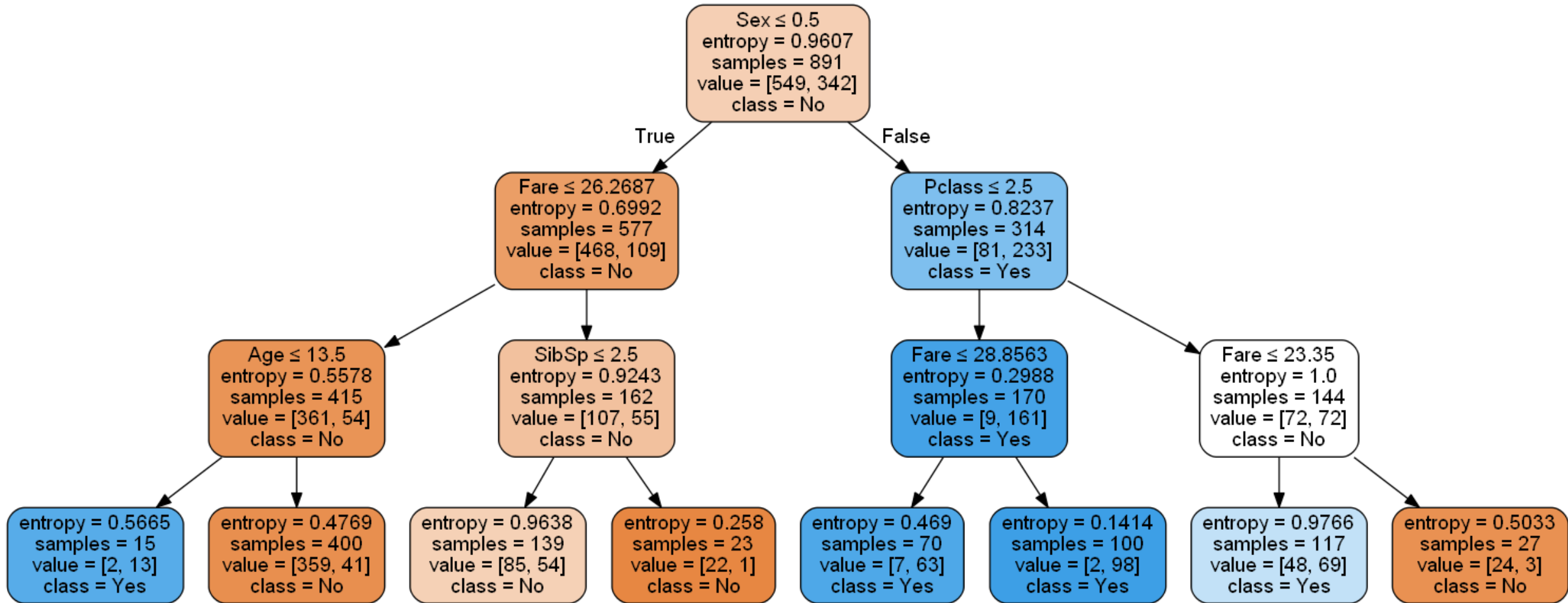
0 = No, 1 = Yes

1 = 1st, 2 = 2nd, 3 = 3rd

C = Cherbourg, Q = Queenstown, S = Southampton

Aquecimento

- Base de dados Titanic



Atividade

- Desenvolver uma das seguintes opções

Árvore de Decisão

Sistema Especialista