

Seleção de Atributos

George Darmiton da Cunha Cavalcanti
gdcc@cin.ufpe.br
CIn - UFPE



UNIVERSIDADE
FEDERAL
DE PERNAMBUCO

CIn.ufpe.br

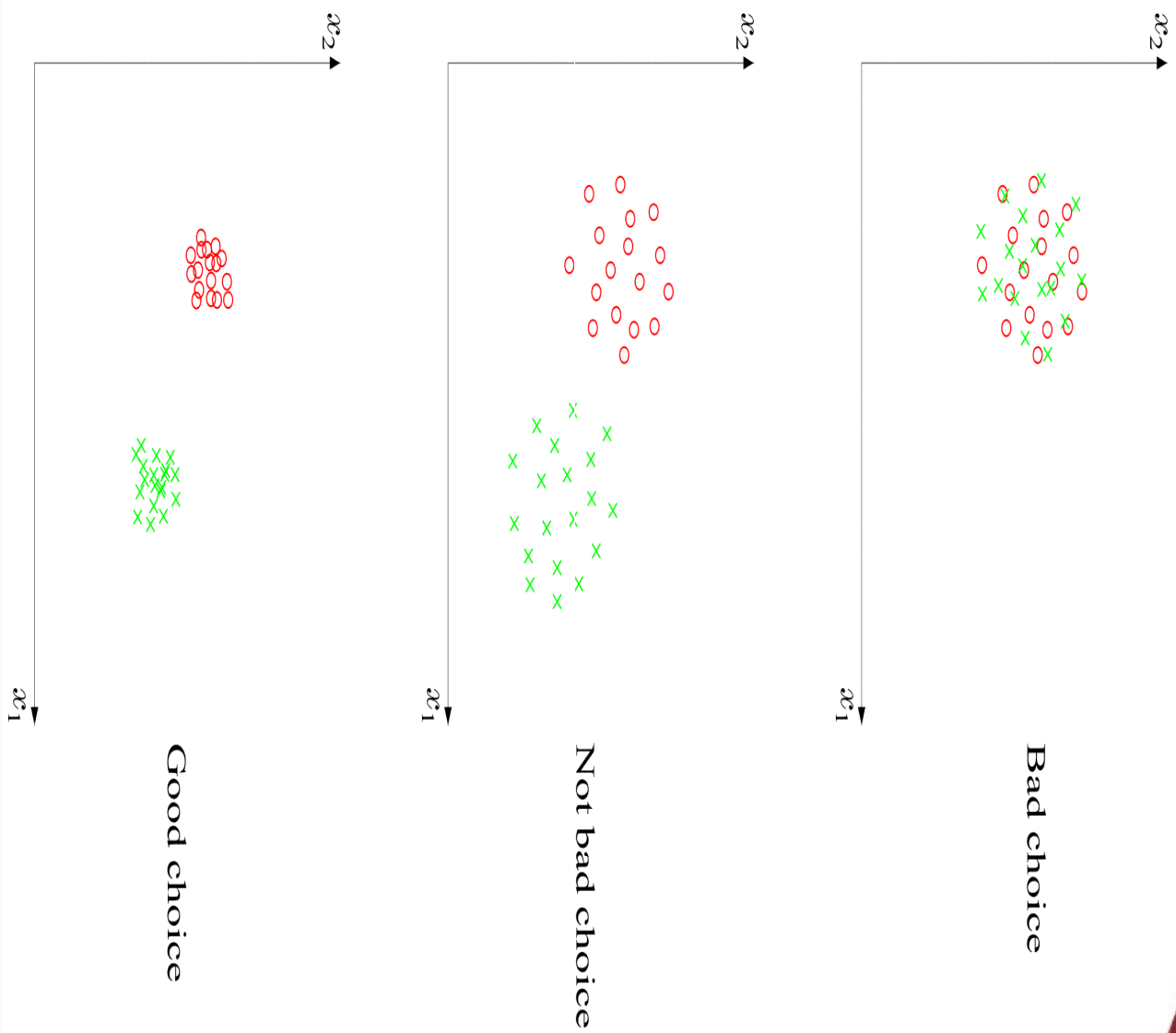
Motivação

- Como **selecionar características** de modo a melhorar a tarefa de classificação?
- O tamanho do conjunto de características afeta diretamente o **design** e o **desempenho** do classificador
- Algumas características não agregam **valor discriminatório**, outras prejudicam o treinamento

Motivação

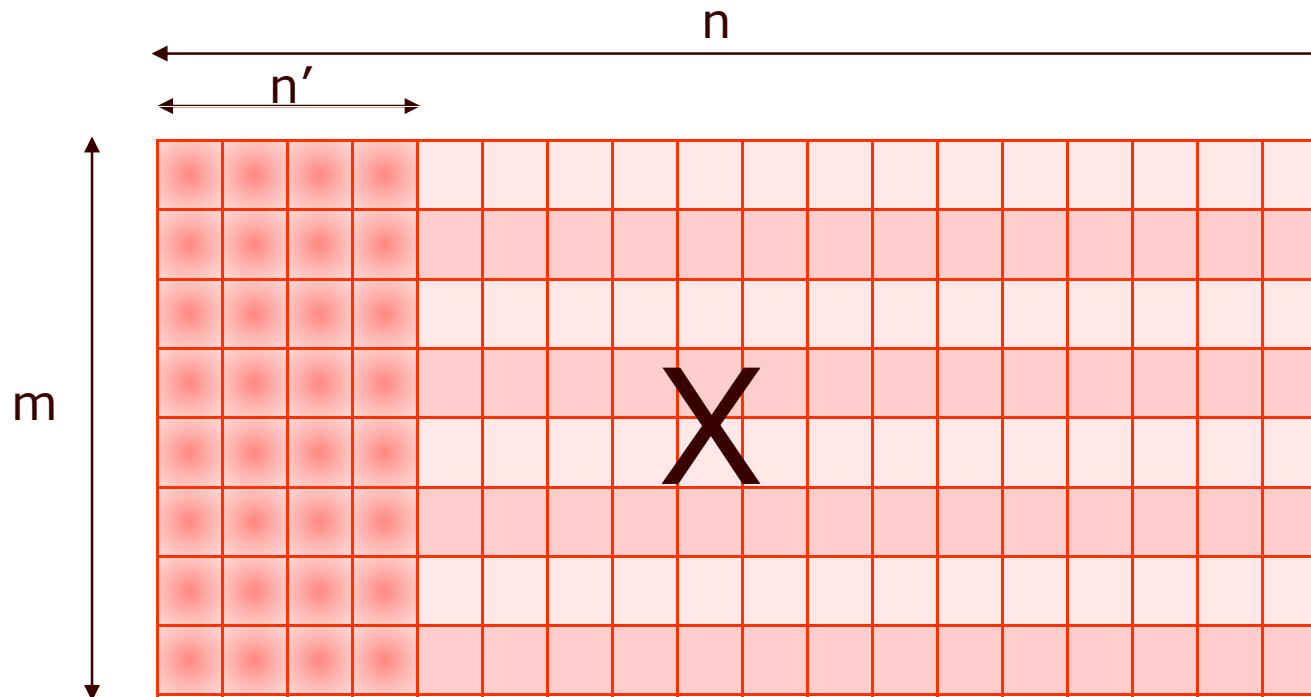
- Quanto maior o número de características...
 - ... maior o **custo computacional** associado
 - ... maior a **difículdade de entendimento** do problema
- É desejado o menor conjunto de características que mantenha a **qualidade da informação discriminatória** entre as classes
- Apesar de sua importância, a escolha das características, muitas vezes, é feita de forma **empírica** ou **subjetiva**

Encontrar os melhores atributos x_1 e x_2



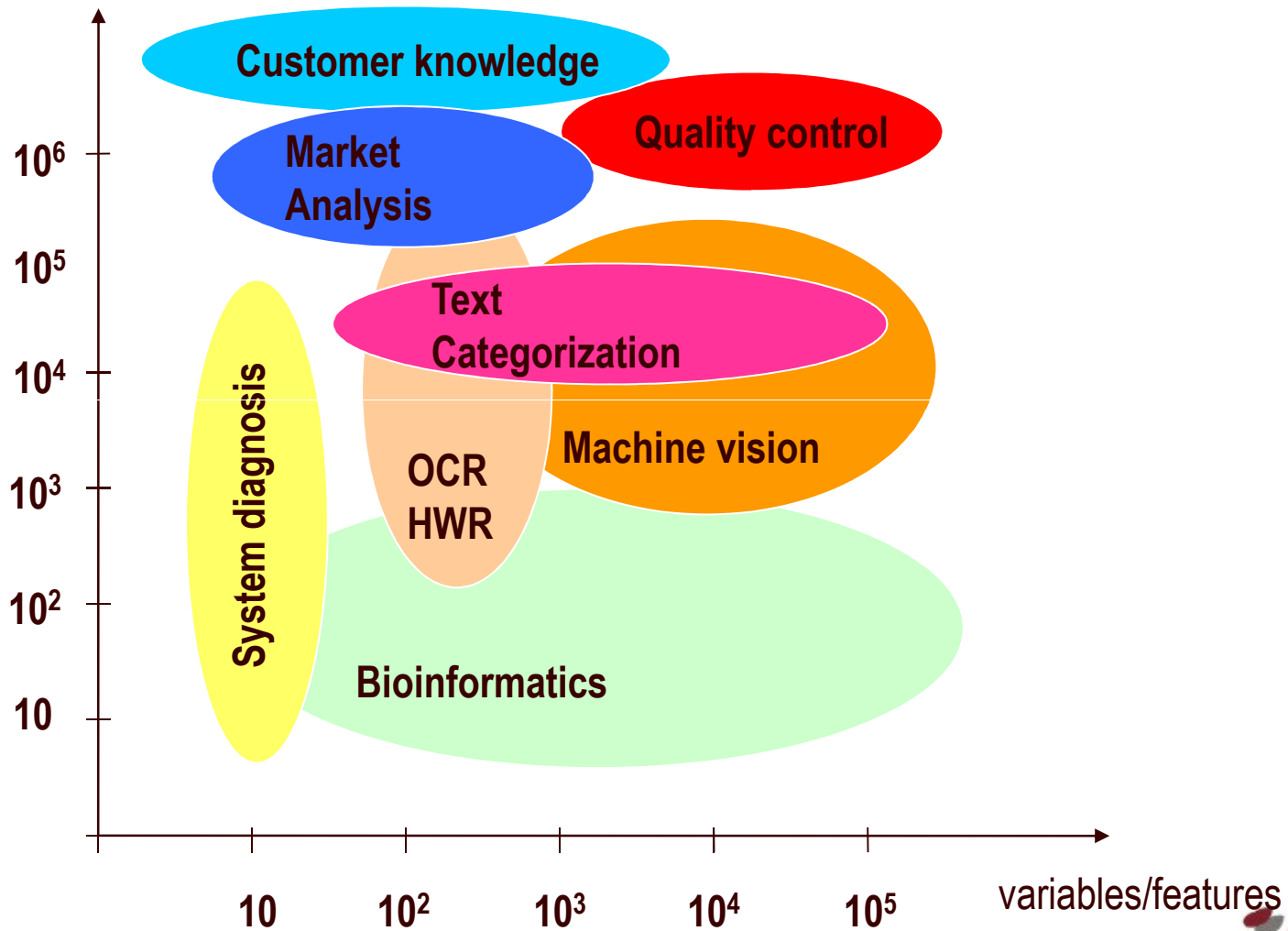
Seleção de Características

- **Das milhares de características:** selecionar as mais relevantes para a construção de máquinas de aprendizagem mais **rápidas, melhores e fáceis de entender**



Aplicações

examples



Seleção de Características

Graus de relevância

$$X = \begin{matrix} & \mathbf{0.1} & \mathbf{0.8} & \mathbf{0.2} & \dots & \mathbf{0.6} \\ \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots & a_{2,m} \\ a_{3,1} & a_{3,2} & a_{3,3} & \dots & a_{3,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \dots & a_{n,m} \end{bmatrix} \end{matrix}$$

Seleção lida com a
redução do número de
características

$m-2$ características

$$X' = \begin{bmatrix} a_{1,2} & a_{1,4} & a_{1,5} & \dots & a_{1,m} \\ a_{2,2} & a_{2,4} & a_{2,5} & \dots & a_{2,m} \\ a_{3,2} & a_{3,4} & a_{3,5} & \dots & a_{3,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,2} & a_{n,4} & a_{n,5} & \dots & a_{n,m} \end{bmatrix}$$

Algoritmos de Seleção: Categorias

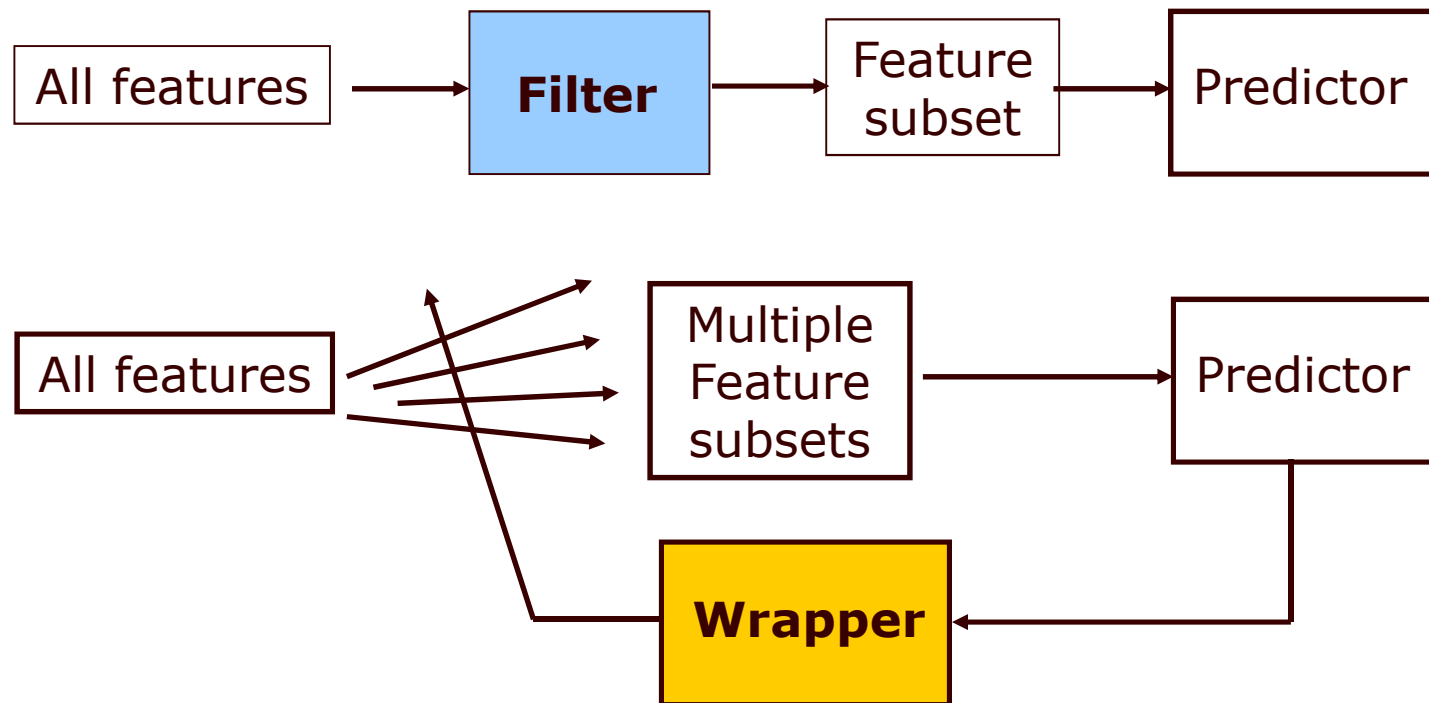
■ Filtros

- *Não faz uso da informação de classificadores*
- *Mais geral*

■ Wrappers

- *Usa a resposta do classificador para selecionar o atributos*
- *Potencialmente específico para o classificador usado*

Filters vs. Wrappers



Árvore de Decisão (Entropia – Ganho de Informação)

- No processo de construção da árvore, o algoritmo ID3 e C4.5, escolhe o “melhor” atributo com base em algum critério
- Objetivo: não selecionar atributos irrelevantes

Possíveis Problemas

■ **Árvore de Decisão**

- Reduzem gradativamente o conjunto de dados usado para realizar a seleção dos melhores atributos

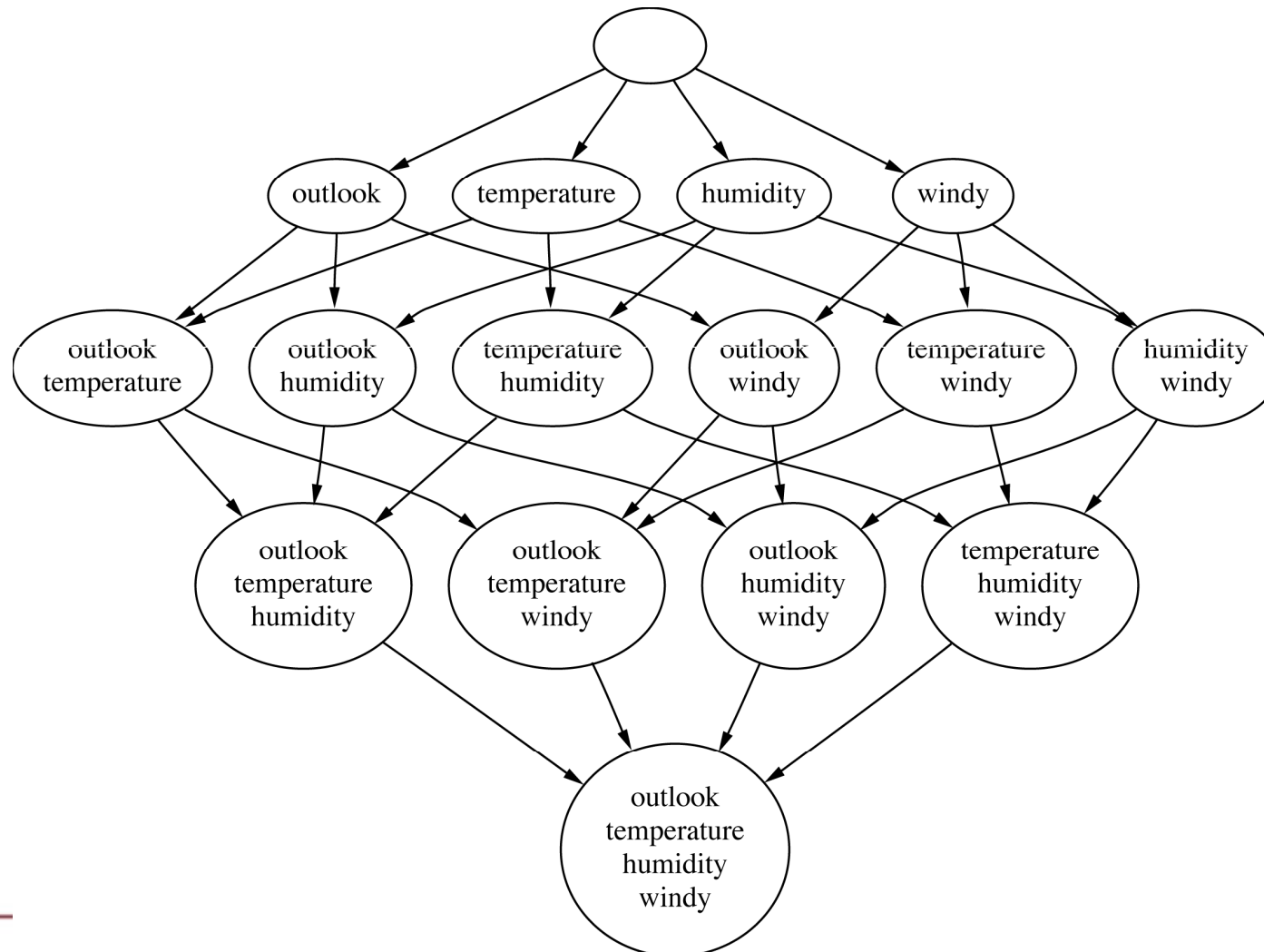
■ **Algoritmos Baseados em Instâncias**

- Observam a vizinhança (algoritmos locais)
- Apenas algumas instâncias são levadas em consideração
- Ex.: k-NN

Tamanho do Espaço de Busca

- Supondo que N é o número total de atributos
- Tem-se, 2^N possíveis subconjuntos
- Assim, uma busca completa (exaustiva) não é computacionalmente viável
- Alternativa: Busca Heurística

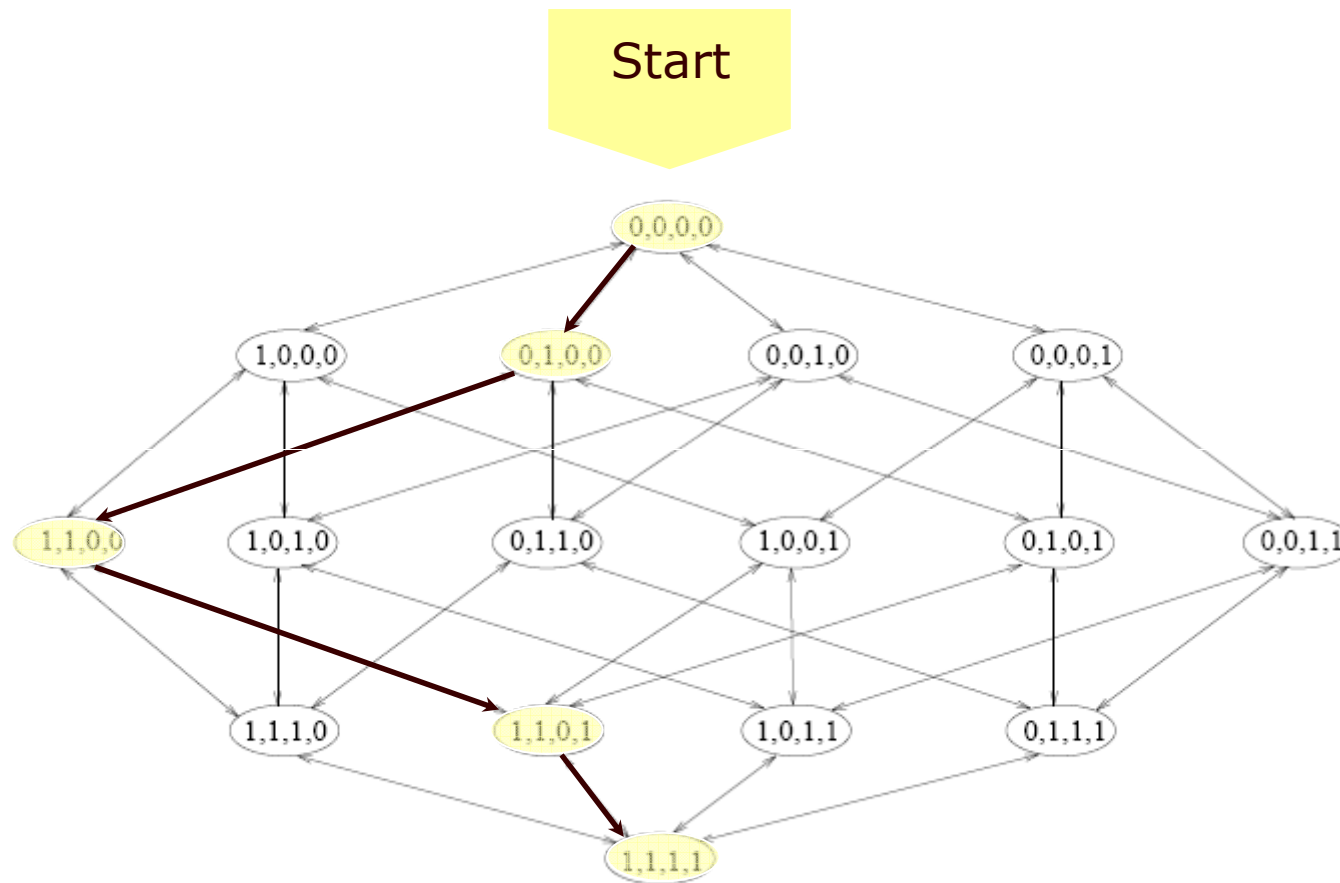
Espaço de Busca (Problema: Jogar Tênis)



Seleção Forward

- A busca é iniciada sem nenhum atributo no conjunto resultante e os atributos são adicionados um a um
- Cada atributo é adicionado isoladamente e o conjunto resultante é avaliado com base em algum critério
- O atributo que produz o “melhor” resultado é adicionado ao conjunto final

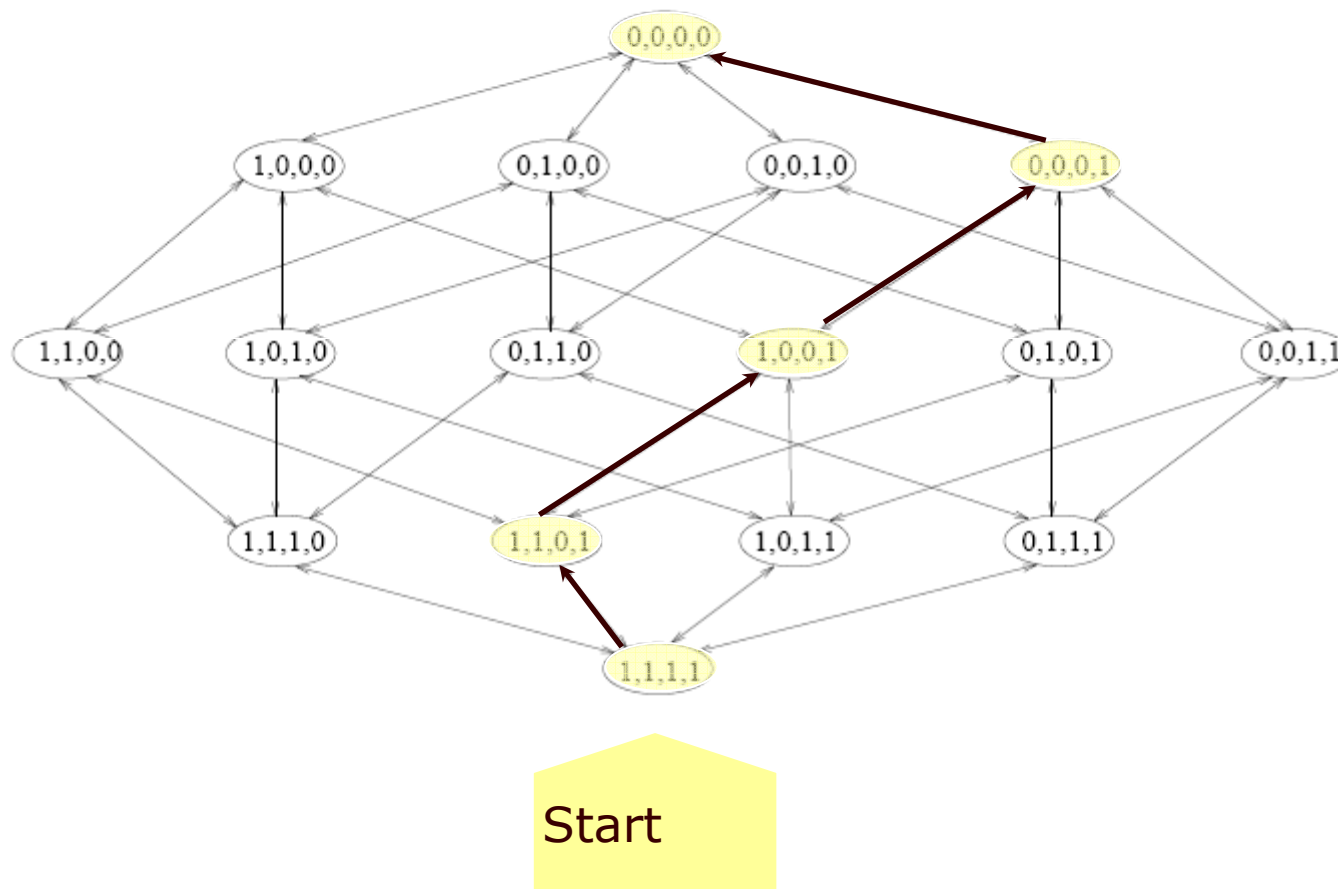
Seleção Forward (Filtro)



Eliminação Backward

- O processo é iniciado com todos os atributos no conjunto final
- A cada passo um atributos é eliminado
- O laço é interrompido com base em algum critério de parada

Eliminação Backward (filtro)



Outras alternativas de algoritmos

- **Busca bidirecional**
- **Best-first search**
- **Algoritmos genéticos**
- **Redes Neurais**
- **Simulated Annealing**
- **Tabu Search**
- **...**

Ponderação de Características

$$X = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,m} \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & a_{3,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & a_{n,m} \end{bmatrix} \quad . * . \quad \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_m \end{bmatrix} \quad \rightarrow$$

$$\rightarrow X' = \begin{bmatrix} w_1 a_{1,1} & w_2 a_{1,2} & w_3 a_{1,3} & \cdots & w_m a_{1,m} \\ w_1 a_{2,1} & w_2 a_{2,2} & w_3 a_{2,3} & \cdots & w_m a_{2,m} \\ w_1 a_{3,1} & w_2 a_{3,2} & w_3 a_{3,3} & \cdots & w_m a_{3,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_1 a_{n,1} & w_2 a_{n,2} & w_3 a_{n,3} & \cdots & w_m a_{n,m} \end{bmatrix}$$

Ponderação preserva o tamanho do conjunto inicial, mas associa pesos a cada característica de acordo com sua relevância