

# Particle swarm optimisation for feature selection e tunning manual

# Objetivos

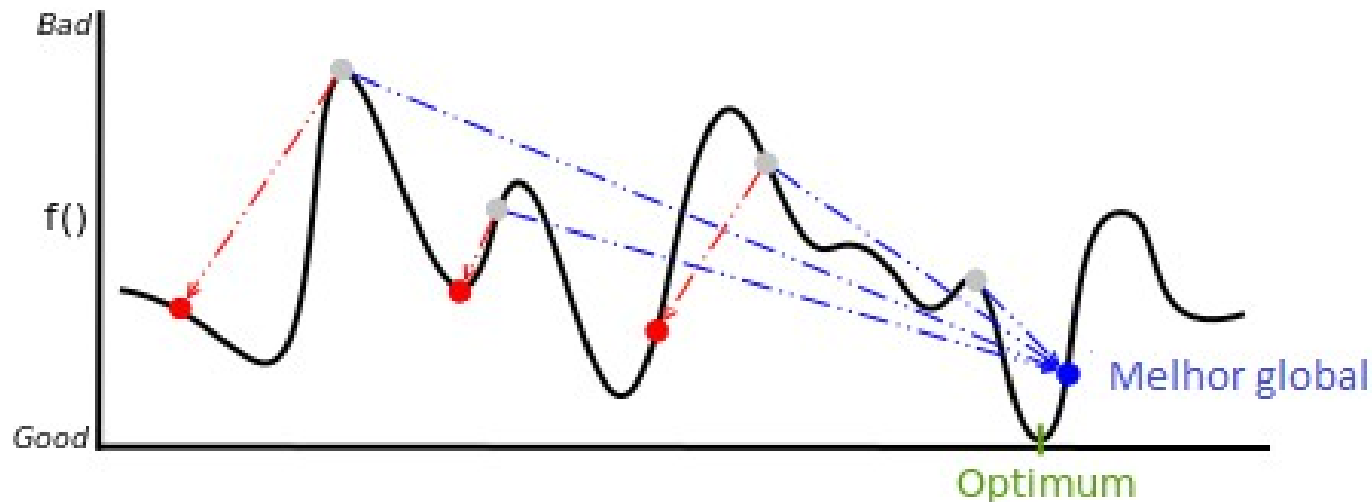
- Apresentar o PSO no contexto de seleção de características
- Apresentar um método de tuning manual dos parâmetros.

# PSO

Particle Swarm Optimization é:

- Baseado em inteligência de enxame
- Approximate
- Não-determinístico

- Durante cada iteração do algoritmo, cada partícula é avaliada por uma função objetivo para avaliar a qualidade do movimento.



# PSO

- Cada partícula contem as seguintes informações:
  - Posição no espaço de busca
  - Velocity
  - Melhor movimento individual
- Adicionalmente, o enxame mantem a melhor a posição global.

# PSO

- O PSO consiste basicamente de três passos:
  - Avaliar cada solução
  - Atualizar os melhores individuais e melhor global
  - Atualizar a velocidade e posição de cada partícula

$$v_i(t+1) = wv_i(t) + c_1r_1[\hat{x}_i(t) - x_i(t)] + c_2r_2[g(t) - x_i(t)]$$

# Estratégias de inicialização

- Três estratégias de inicialização baseados em **small initialisation**(forward selection), **large initialisation**(backward selection) e **mixed initialization**(com as vantagens das duas técnicas)
- **Small initialisation**: Cada partícula é inicializada com uma pequena porcentagem(10%) do total de features e de maneira aleatória.
- **Large initialisation**: Cada partícula é inicializada usando uma grande quantidade de features(entre metade + 1 e total de features) e de maneira aleatória
- **Mixed Initialisation**: 2/3 das partículas é inicializada com a estratégia small initialisation e 1/3 inicializado com a estratégia large initialisation

# Estratégias de atualização

## 1) Primeira estratégia e Segunda estratégia

```

if accuracy of particle  $i$  ( $x_i$ ) is better than  $pbest$  then
    |  $pbest = x_i$  ; // Update the  $pbest$ 
else if accuracy of  $x_i$  is the same as  $pbest$  and  $|x_i| < |pbest|$  then
    |  $pbest = x_i$  ; // Update the  $pbest$ 
if accuracy of any  $pbest$  is better than  $gbest$  then
    |  $gbest = pbest$  ; // Update the  $gbest$ 
else if accuracy of any  $pbest$  is the same as  $gbest$  and  $|pbest| < |gbest|$  then
    |  $gbest = pbest$  ; // Update the  $gbest$ 

```

```

if accuracy of particle  $i$  ( $x_i$ ) is better than pbest then
    |   pbest =  $x_i$  ;                               // Update the pbest
if accuracy of any pbest is better than gbest then
    |   gbest = pbest ;                               // Update the gbest

```

# Tunning Manual

- Selecione os parâmetros mais influentes.
- Estratégia
  - Underfitting
  - Good fit
  - Overfitting
- Dois grupos:
  - Increasing it impedes fitting ( overfitting to underfitting)  
min\_child\_weight, lambda, alpha
  - Increasing it leads to a better fit ( underfitting to overfitting)  
max\_depth, subsample, colsample\_bytree