

Machine Learning para previsão de Séries Temporais

Prof. Eraylson Galdino



Agenda

- Motivação
- Técnicas de ML
- Etapas
 - Pré-processamento
 - Treinamento
 - Avaliação
- Sistemas Híbridos
- Ensemble



ML para séries temporais

- Pontos positivos:
 - Modelos Não lineares;
 - Aproximadores universal de funções;
 - Não paramétricos;
- Pontos negativos:
 - Overfitting
 - Underfitting
 - Seleção de Hiperparâmetros



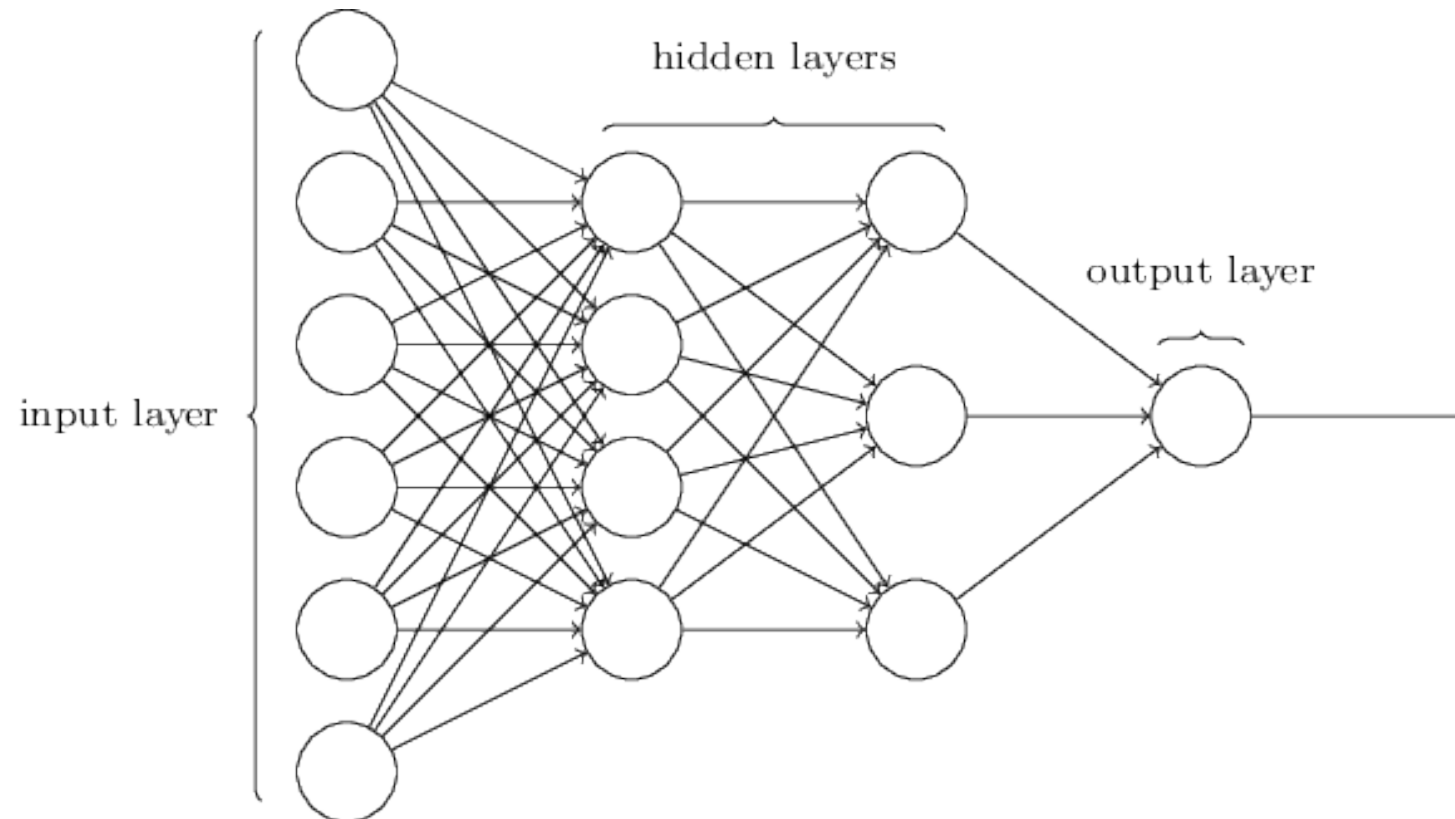
ML para séries temporais

- Principais técnicas:
 - Redes Neurais MLP
 - Extreme Learning Machine
 - Rede de Base Radial
 - Máquina de Vetor de Suporte
 - Long Short-Term Memory



ML para séries temporais

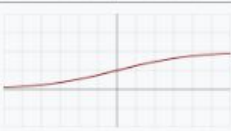
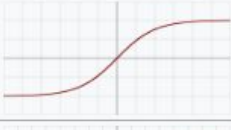

- Redes Neurais MLP





ML para séries temporais

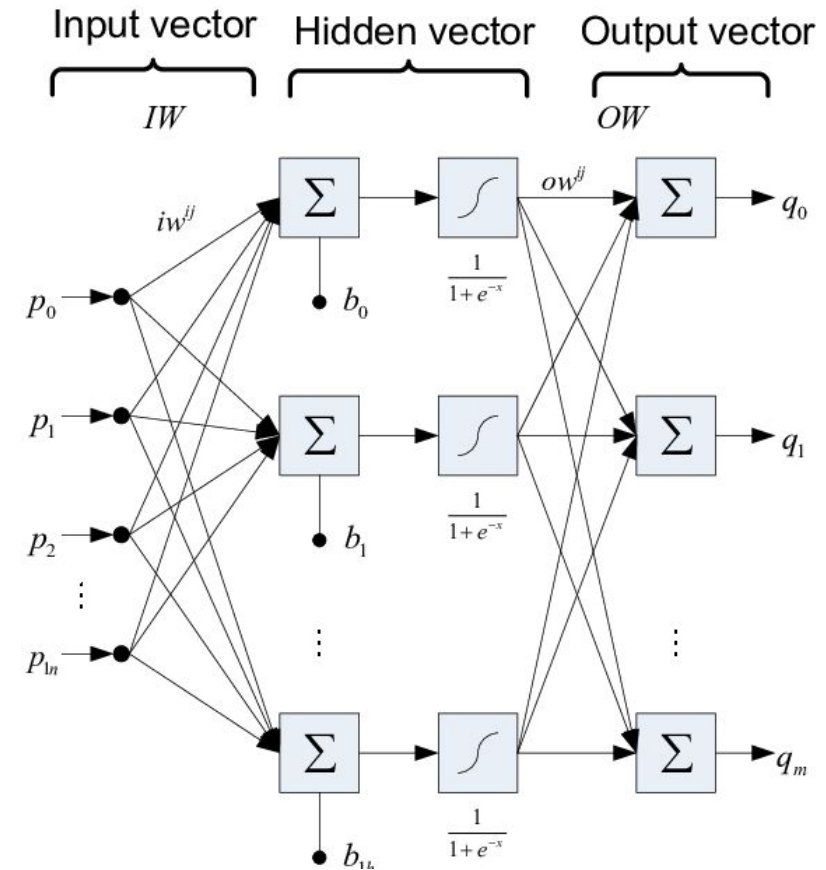
- Redes Neurais MLP (funções de ativação)

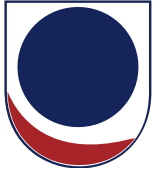
Logistic (a.k.a. Sigmoid or Soft step)		$f(x) = \sigma(x) = \frac{1}{1 + e^{-x}}$ ^[1]	$f'(x) = f(x)(1 - f(x))$	$(0, 1)$
TanH		$f(x) = \tanh(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})}$	$f'(x) = 1 - f(x)^2$	$(-1, 1)$
Rectified linear unit (ReLU) ^[14]		$f(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ x & \text{for } x > 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ 1 & \text{for } x > 0 \end{cases}$	$[0, \infty)$



ML para séries temporais

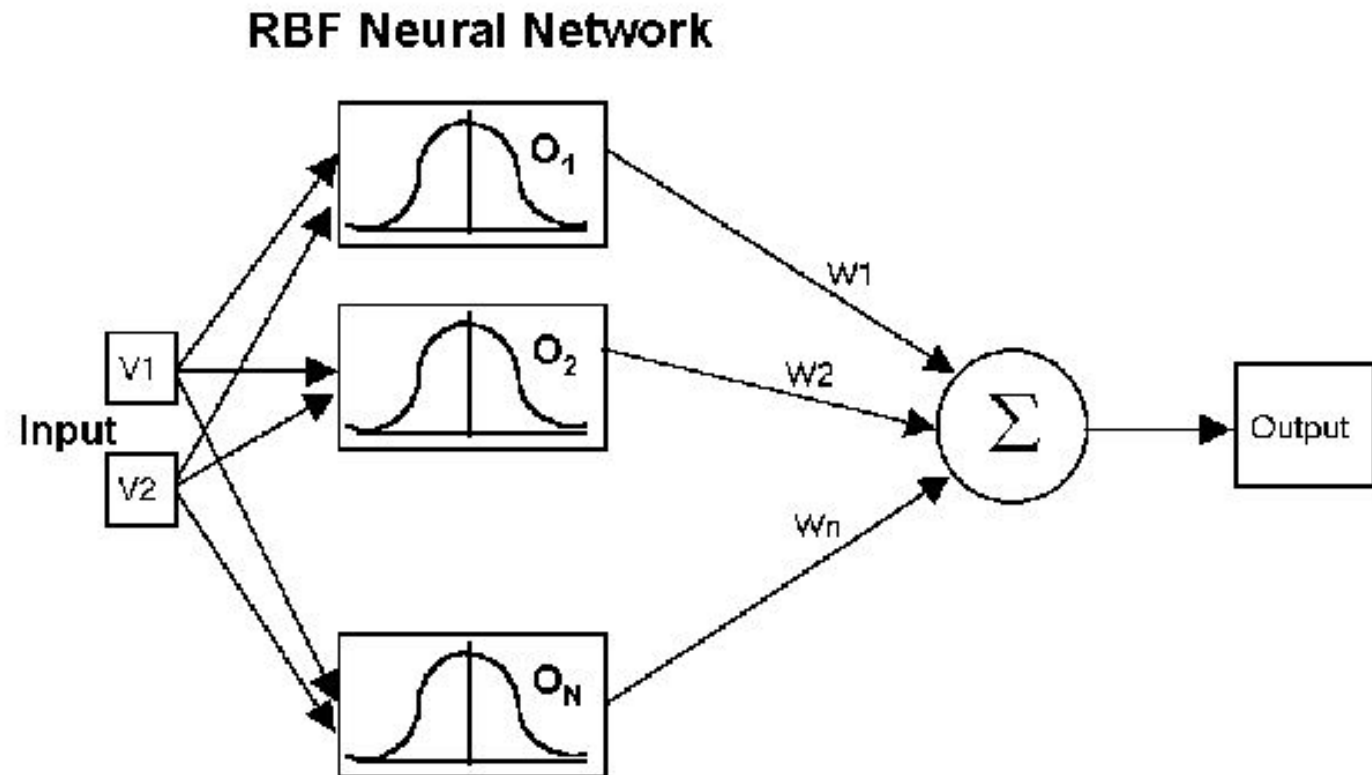
- Extreme Learning Machine





ML para séries temporais

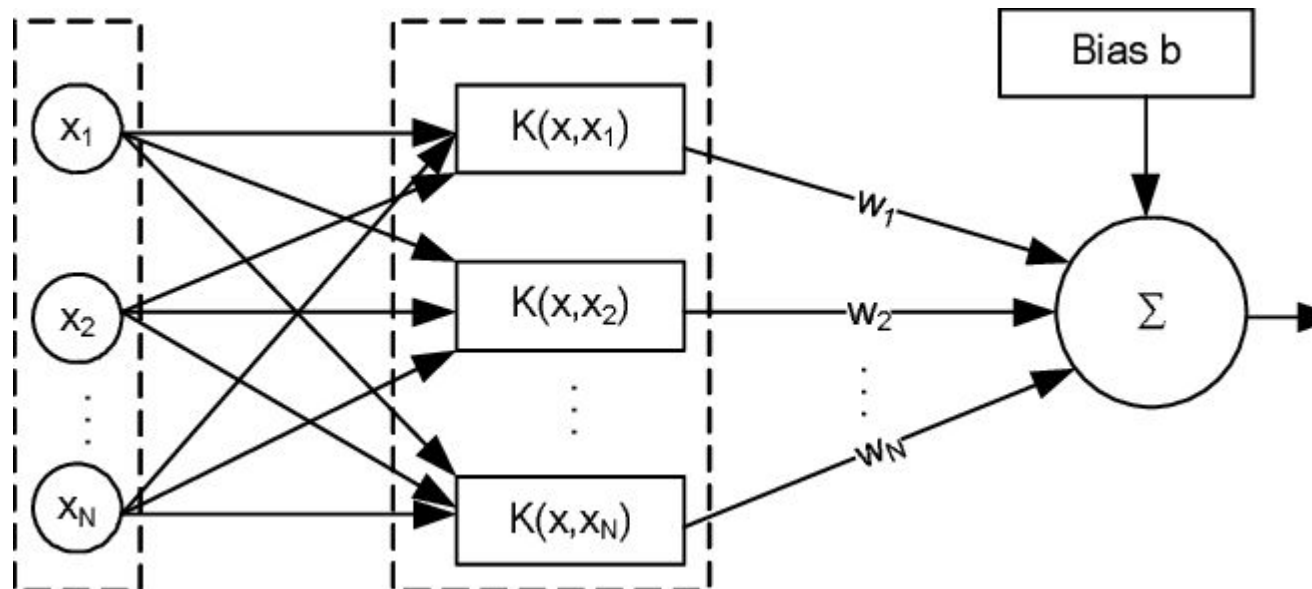
- Rede de base Radial





ML para séries temporais

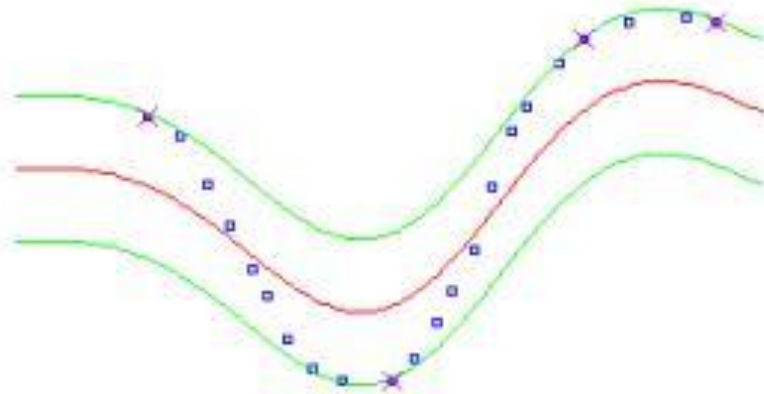
- Máquina de Vetor de Suporte





ML para séries temporais

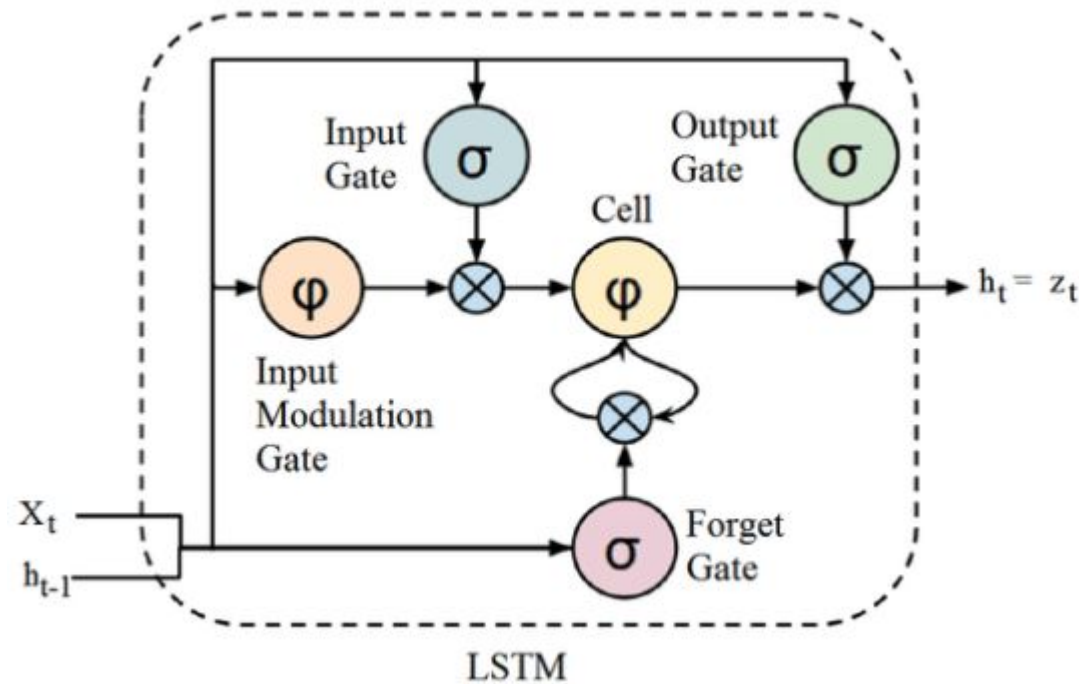
- Máquina de Vetor de Suporte

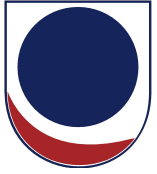




ML para séries temporais

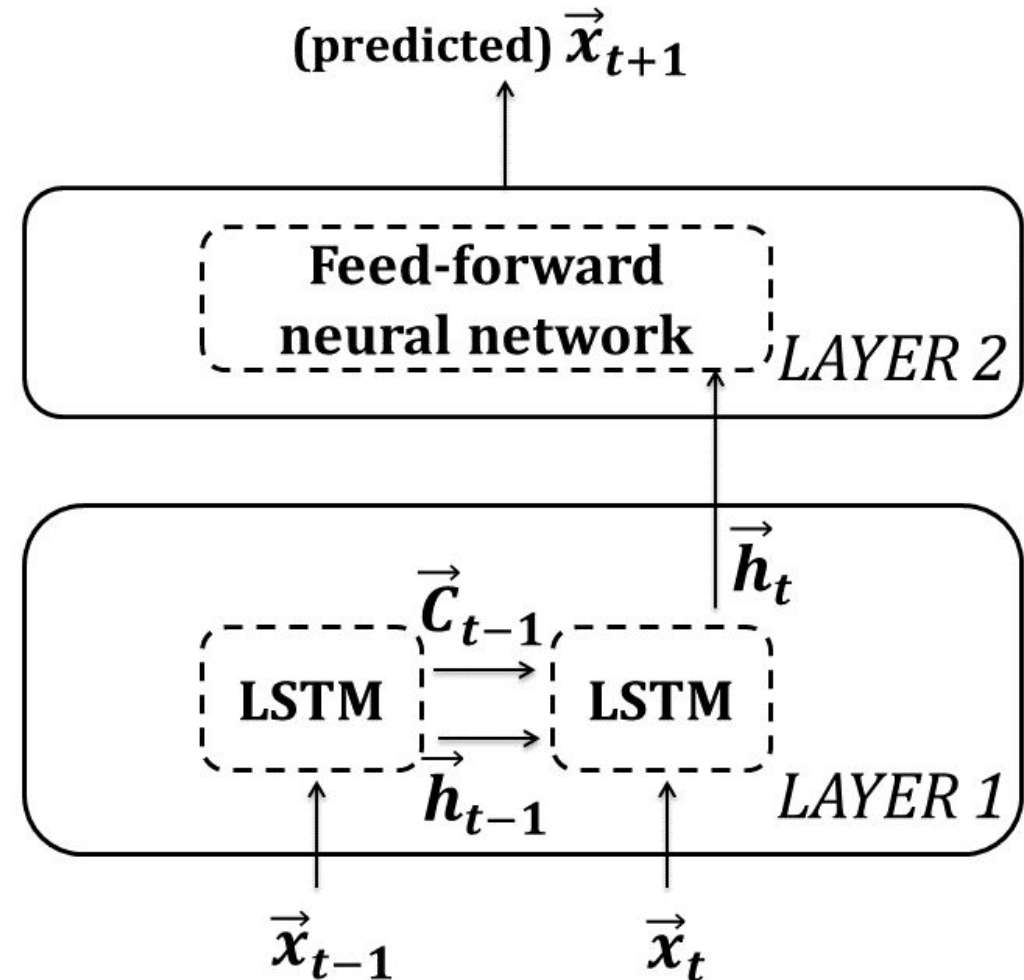
- Long Short-Term Memory





ML para séries temporais

- Long Short-Term Memory





ML aplicado para previsão

Pré-processamento

- Diferenciação
- Normalização
- Transformação
- Janelamento
- Particionamento

Treinamento

- Gridsearch
- algoritmo de otimização

Avaliação

- MAPE
- MSE
- POCID

Aplicação

- Deploy



Pré-processamento

- Diferenciação:
 - Aplicado para tornar a série estacionária;
 - Indicado aplicar antes de treinar modelos com problemas de desempenho em séries com tendência e sazonalidade;



Pré-processamento

- Normalização:
 - Transformar os dados para uma escala adequada para os modelos;

$$Z'(t) = \frac{Z(t) - \min(Z)}{\max(Z) - \min(Z)} \longrightarrow [0, 1]$$

$$Z'(t) = \frac{Z(t) - \max(Z) - \min(Z)}{\max(Z) - \min(Z)} \longrightarrow [-1, 1]$$

$$Z'(t) = \frac{Z(t) - \text{média}(Z)}{\text{desvio}(Z)} \longrightarrow N(0, 1)$$



Pré-processamento

- Transformação em log:
 - Aplicado em séries com heteroscedasticidade
 - $Z' = \log(Z)$



Pré-processamento

- Janelamento:
 - Organiza a série para ser utilizada em modelos de aprendizagem supervisionada

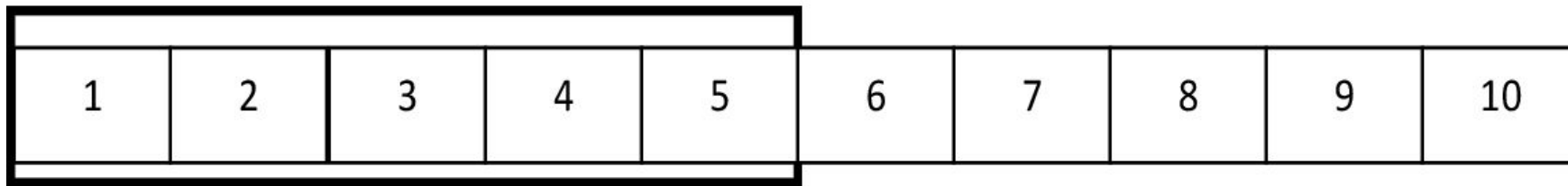
X1	X2	X3	Y
1	2	3	4
2	3	4	5
3	4	5	6



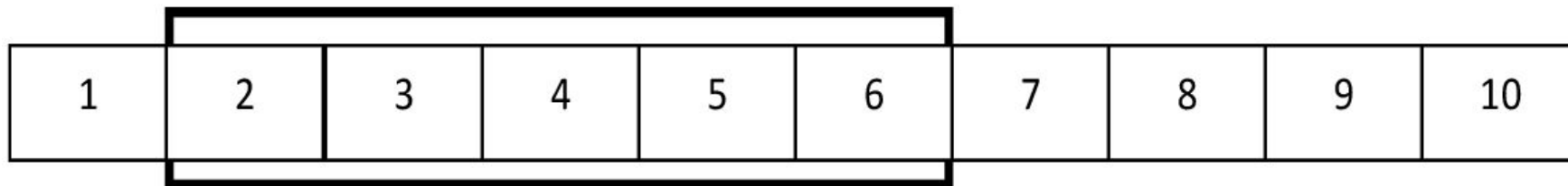
Pré-processamento

- Janelamento:

Initial Window



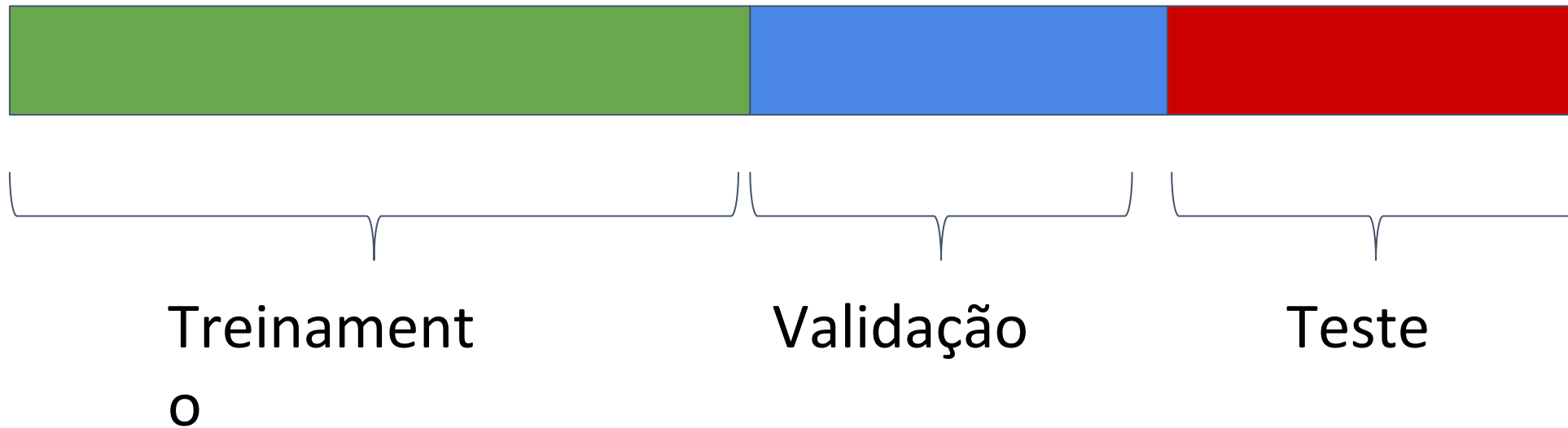
Window Slide

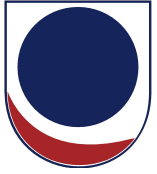




Pré-processamento

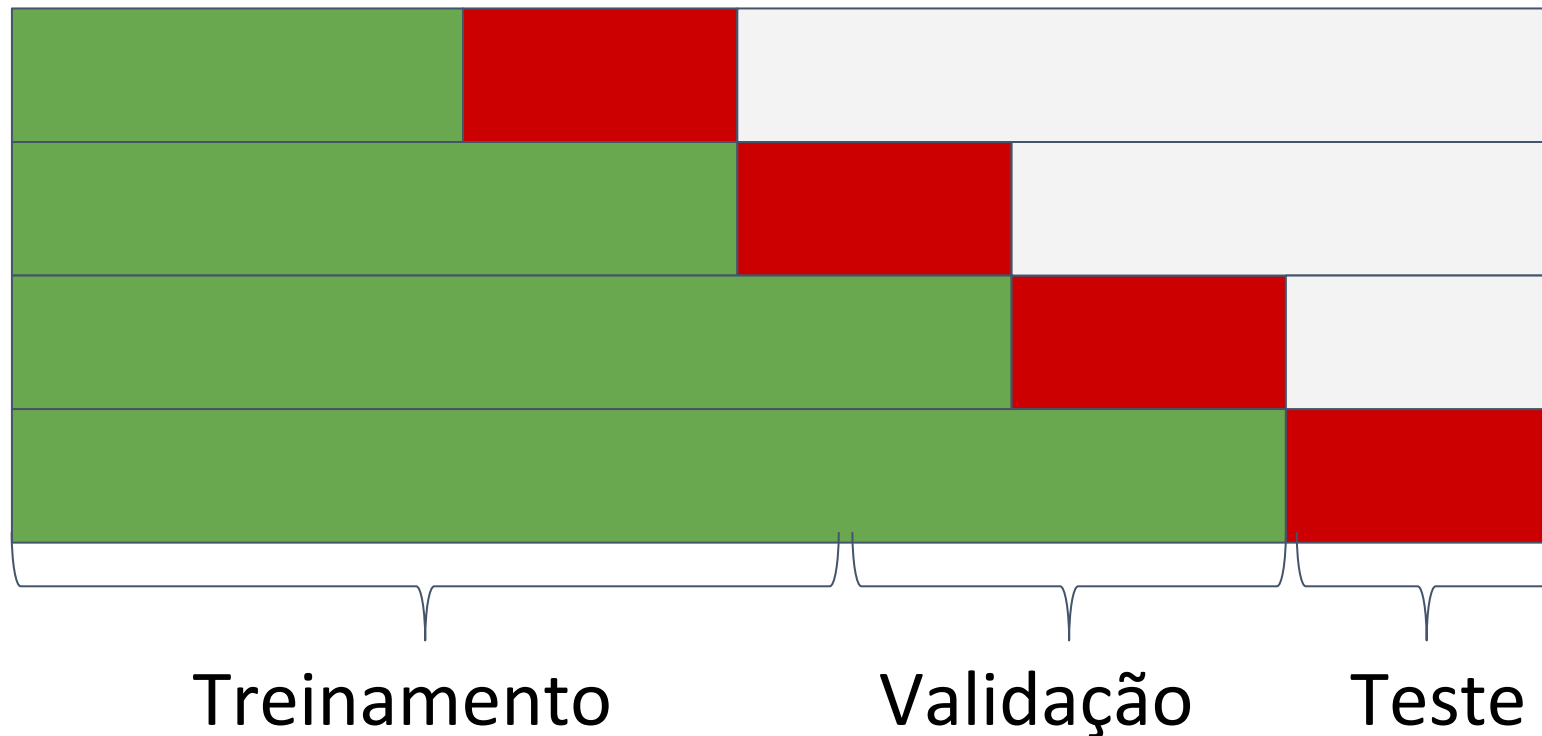
- Particionamento: (Hold-out)





Pré-processamento

- Particionamento: (Cross-validation)





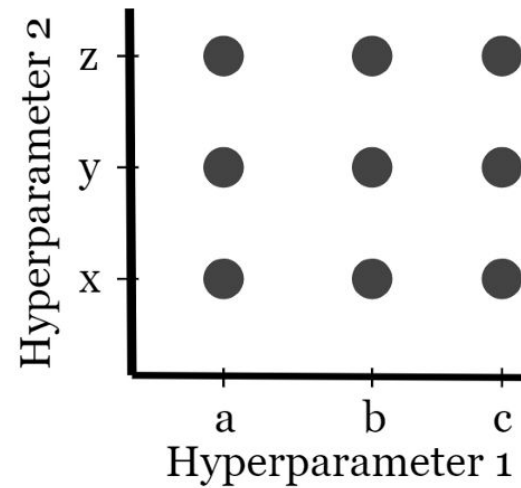
Treinamento

- Gridsearch

Grid Search

Pseudocode

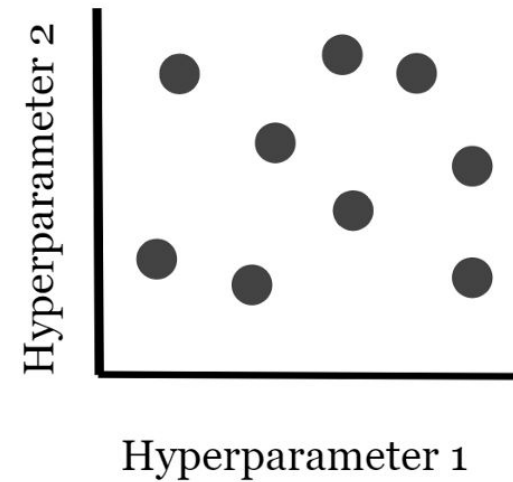
```
Hyperparameter_One = [a, b, c]  
Hyperparameter_Two = [x, y, z]
```



Random Search

Pseudocode

```
Hyperparameter_One = random.num(range)  
Hyperparameter_Two = random.num(range)
```





Treinamento

- Algoritmos de otimização bio-inspirados
 - Algoritmos Evolucionários:
 - Algoritmo Genético

cromossomo = [hp_1, hp_2, ..., hp_n]

- Inteligência de Enxames:
 - PSO

partícula = [hp_1, hp_2, ..., hp_n]



Aviação

- Medidas de desempenho

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{desejado(i) - previsto(i)}{desejado(i)} \right|,$$

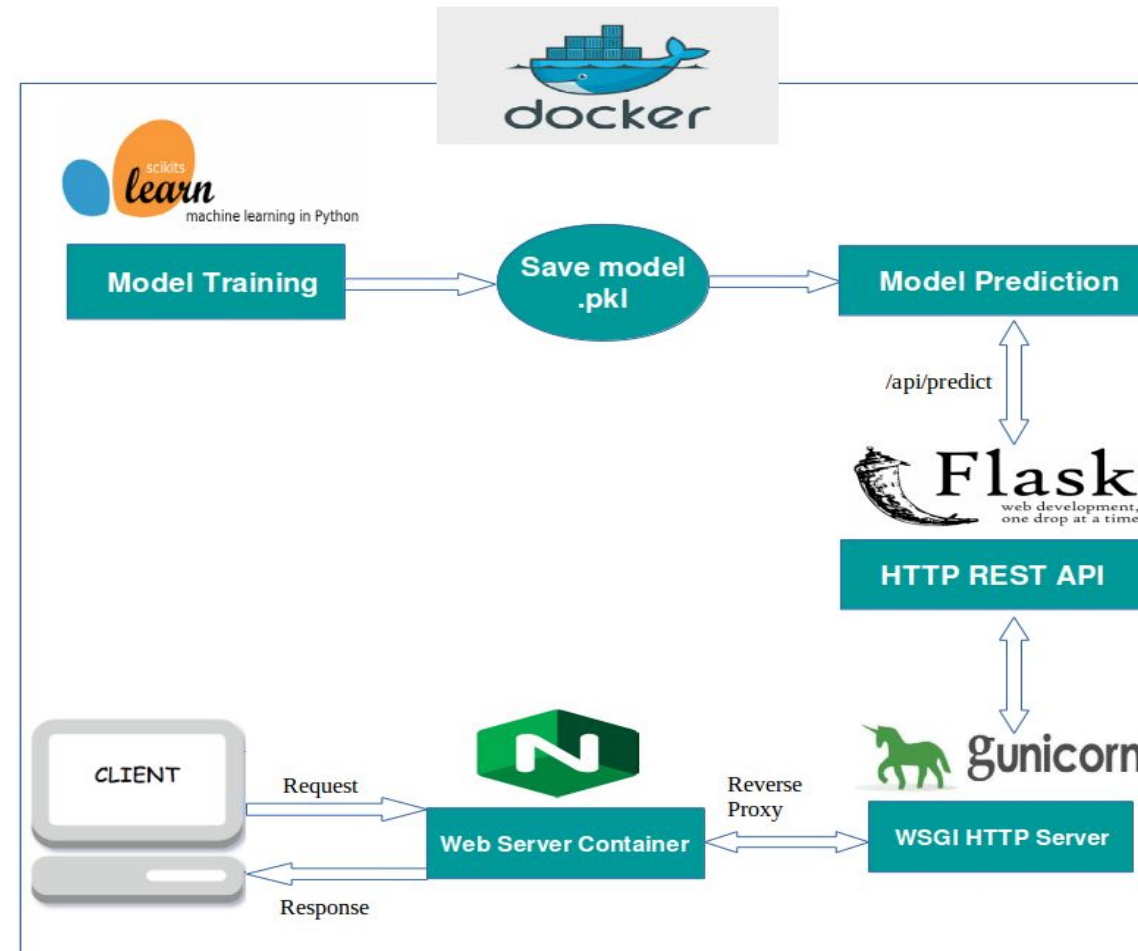
$$MSE = \frac{1}{N} \sum_{i=1}^N (desejado(i) - predito(i))^2$$

$$POCID = \frac{100}{N} \sum_{i=1}^N X_i,$$

$$X_i = \begin{cases} 1, & \text{se } (desejado(t) - desejado(t-1))(previsto(i) - previsto(t-1)) > 0 \\ 0, & \text{caso contrário} \end{cases}$$



Aplicação



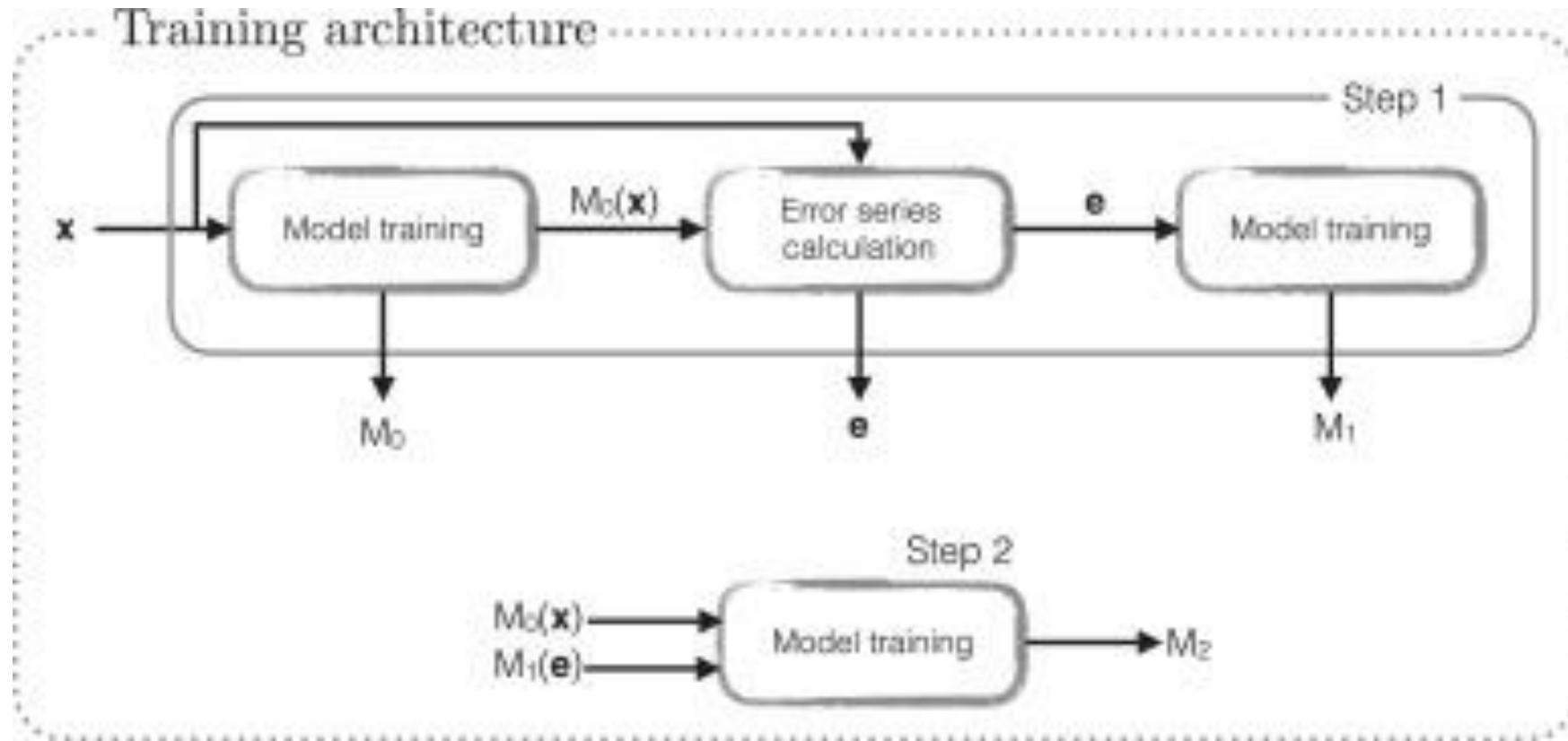


Sistemas Híbridos

- Modelos estatísticos são bons para séries lineares;
- Modelos de AM são bons para séries não lineares;
- Séries do mundo real tem comportamento linear e não linear;
- Composição de modelos estatísticos com modelos de AM.



Sistemas Híbridos





Sistemas de Múltiplos Preditores

- Motivação:
 - Teorema do NFL;
 - Não existe um conjunto de regras para a escolha do melhor preditor;
 - O melhor preditor para o conjunto de dados atual pode ser o pior para dados futuros;



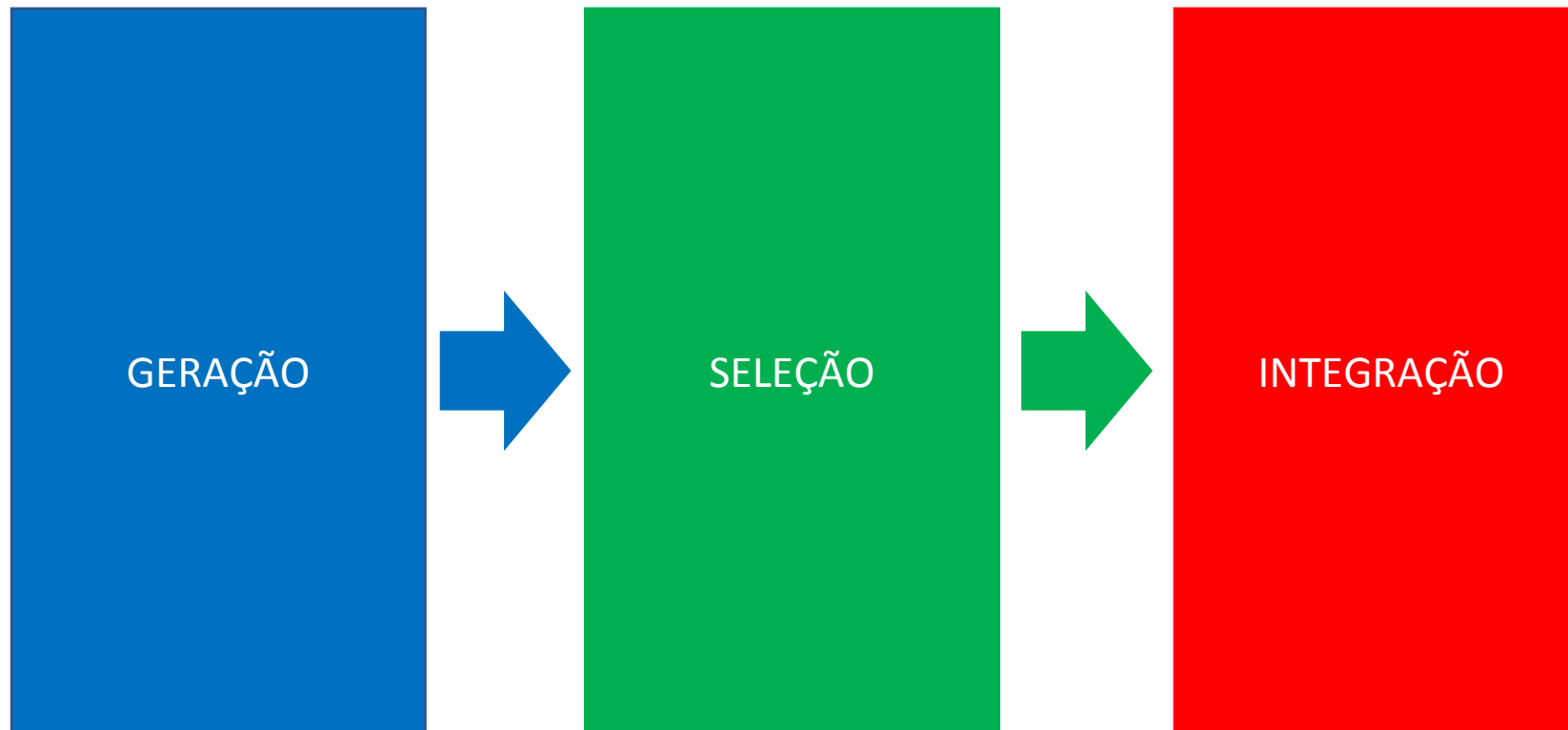
Sistemas de Múltiplos Preditores

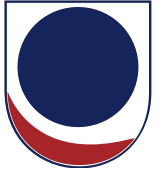
- Definição:
 - Pode ser definido como um conjunto de modelos diferentes com o objetivo de prever a mesma variável
 - A previsão final de um SMP é realizada com base na combinação de todas as previsões obtidas ou através da seleção de uma previsão;



Sistemas de Múltiplos Preditores

- Etapas





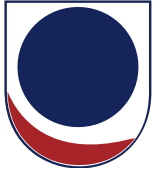
Sistemas de Múltiplos Preditores

- Geração:
 - Consiste em criar o conjunto de modelos (Pool);
 - É necessário que exista diversidade entre os modelos:
 - Modelos treinados através de diferentes amostras;
 - Bagging, Split, AdaBoost, EMD, Crogging;
 - Modelos utilizando diferentes parâmetros;
 - Diferentes Modelos;



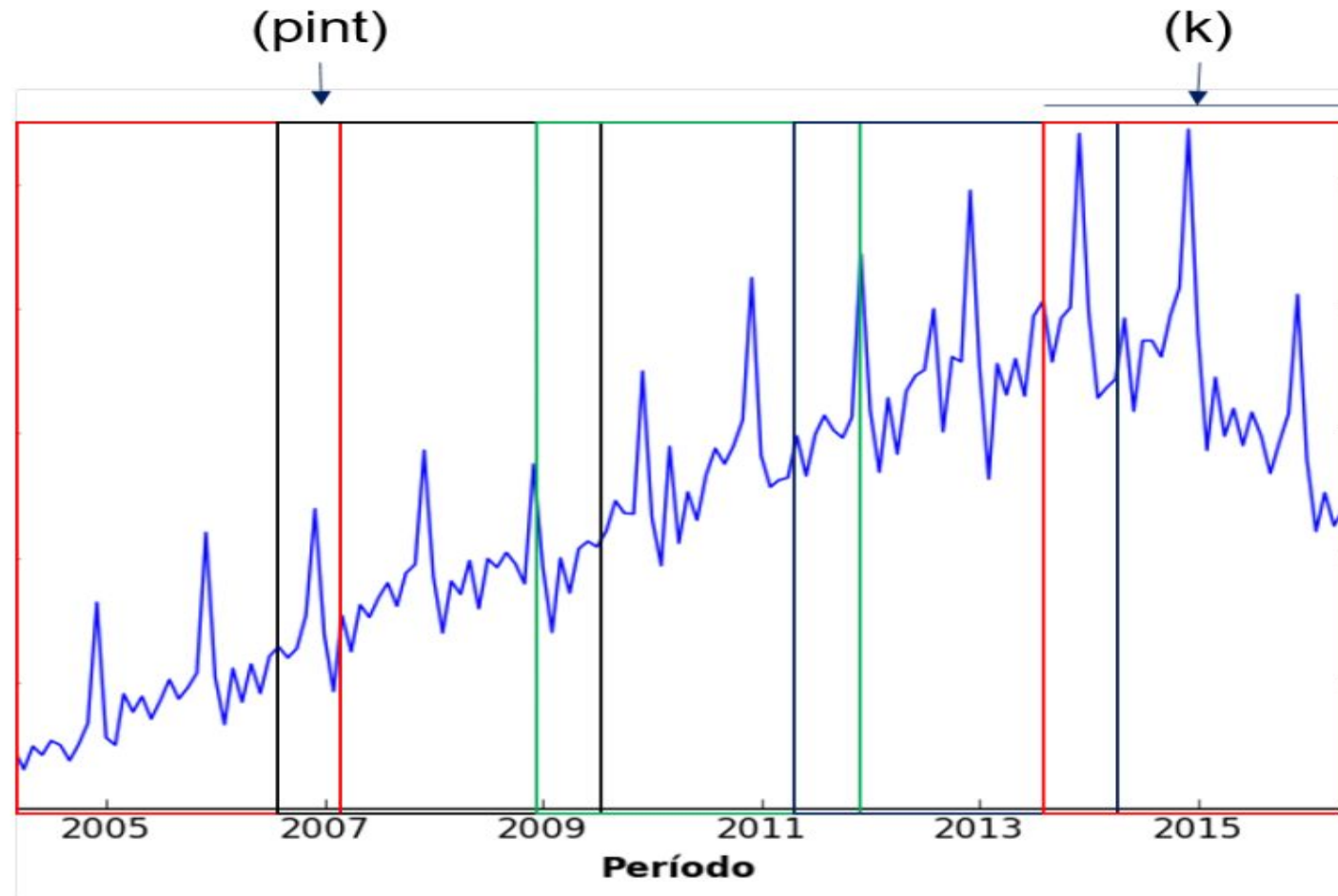
Sistemas de Múltiplos Preditores

- Geração
 - Bagging
 - `T <- Conjunto de treinamento`
 - `Pool <- []`
 - **De 1 até Qtd_de_preditores Faça:**
 - `S <- reamostragem(T)`
 - `P <- treina_preditor(S)`
 - `Pool <- Pool + P`
 - **Retorna Pool**



Sistemas de Múltiplos Preditores

- Split





Sistemas de Múltiplos Preditores

- Seleção
 - Estática: Consiste em selecionar um ou mais preditores do pool para realizar a previsão para todos os novos dados;
 - Geralmente os preditores são selecionados com base no desempenho obtido no conjunto de validação;



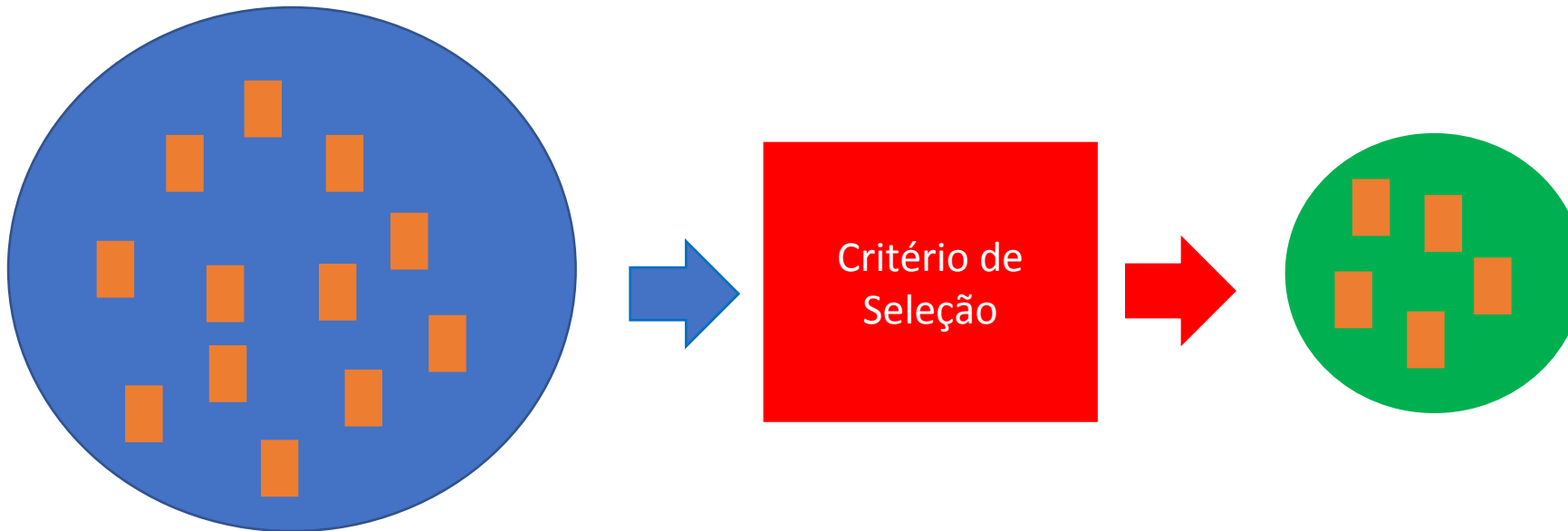
Sistemas de Múltiplos Preditores

- Seleção
 - preditores do pool para cada novo dado;
 - Preditores são selecionados de acordo com o novo padrão. Geralmente é utilizado a similaridade entre o novo padrão e os padrões anteriores, ou é utilizado de Metalearning;
 - Oracle: Desempenho do melhor Preditor para o ponto (t);

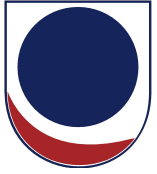


Sistemas de Múltiplos Preditores

- Seleção

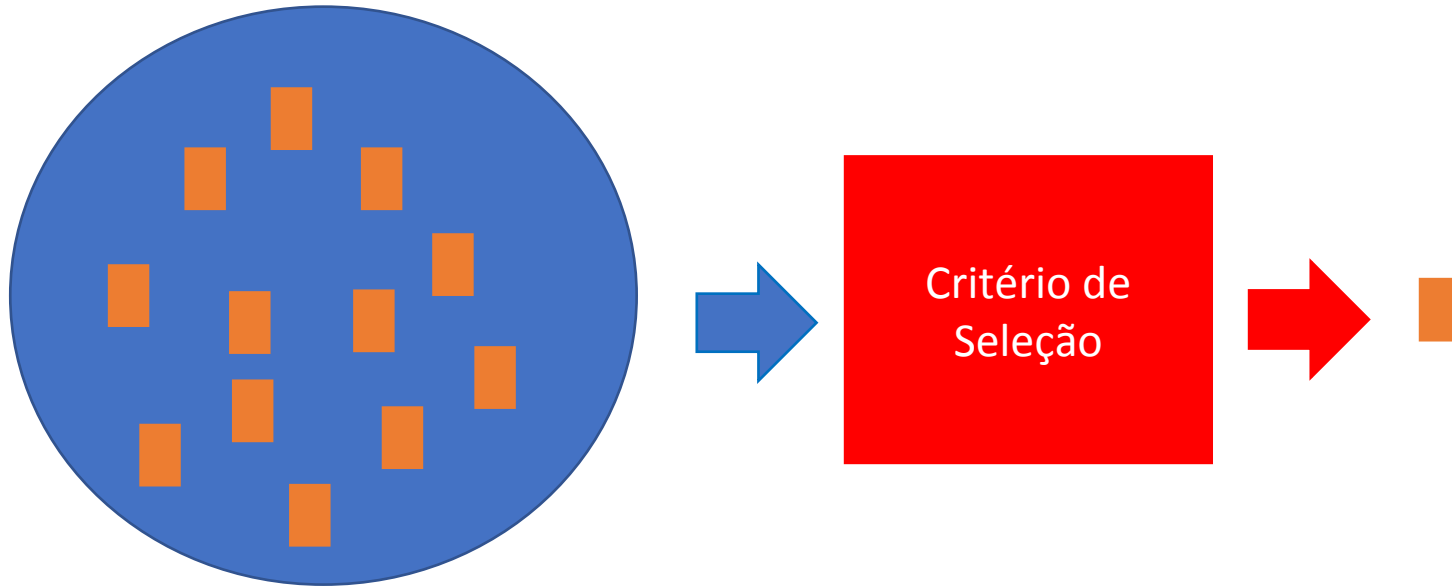


Seleção estática de Ensemble



Sistemas de Múltiplos Preditores

- Seleção

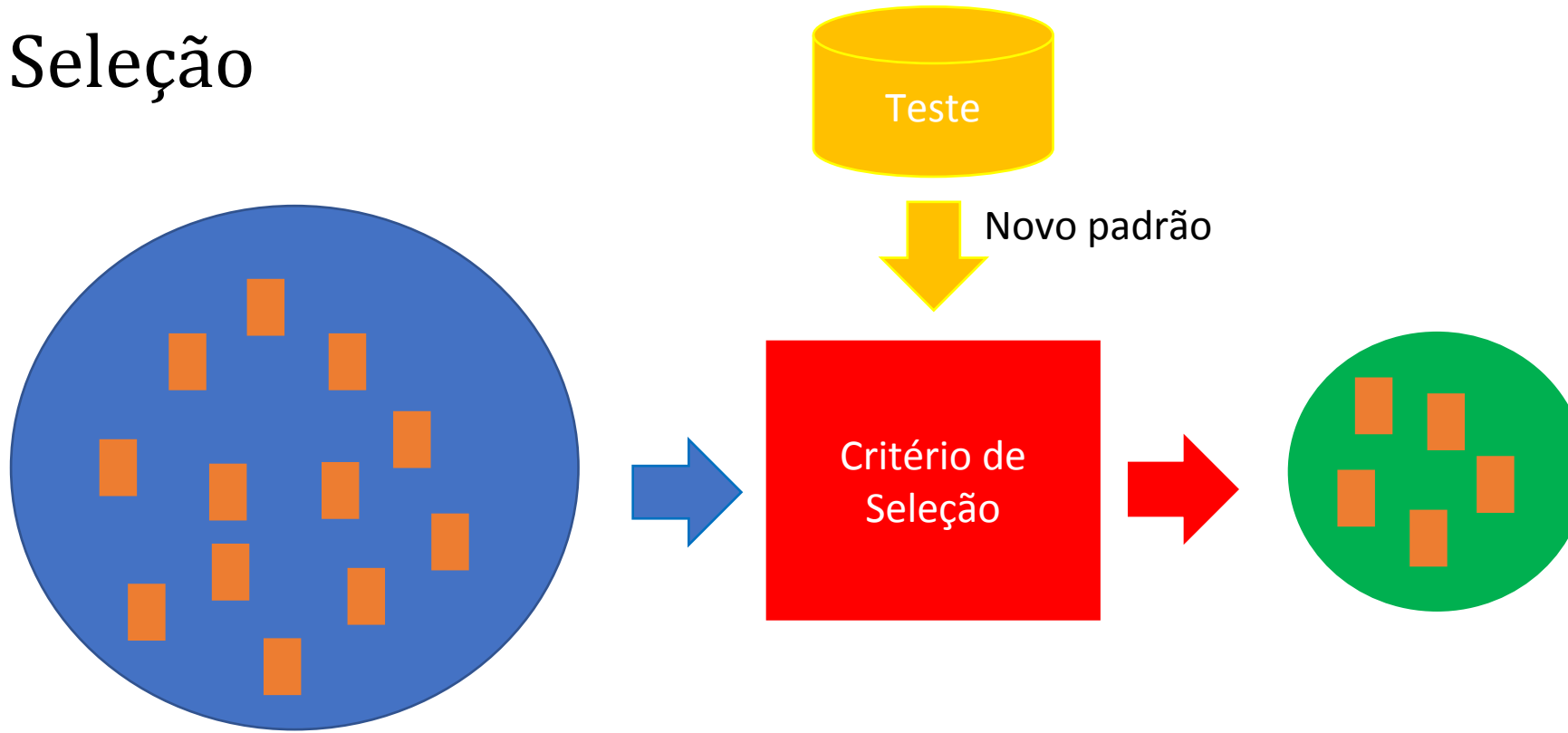


Seleção estática de Preditor



Sistemas de Múltiplos Preditores

- Seleção

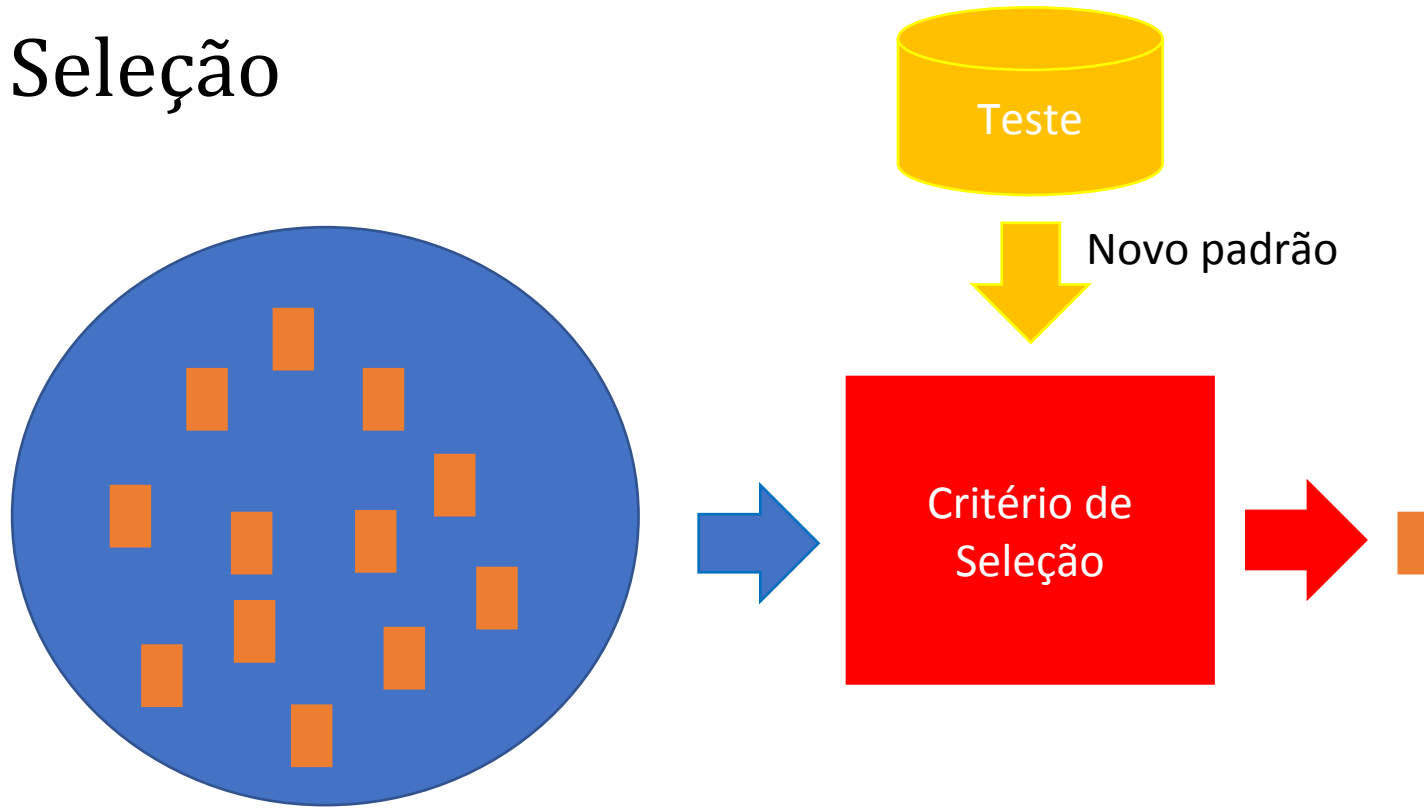


Seleção dinâmica de Ensemble



Sistemas de Múltiplos Preditores

- Seleção

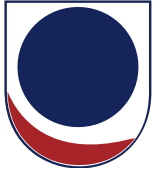


Seleção dinâmica de Preditor



Sistemas de Múltiplos Preditores

- Integração:
 - Consiste em integrar as previsões dos modelos em apenas uma previsão final;
 - A seleção dinâmica/estática de um preditor pode ser considerada uma forma de integração;
 - Nessa etapa é realizado a combinação das previsões do Ensemble;



Sistemas de Múltiplos Preditores

- Integração:
 - Possíveis técnicas:
 - Linear Ponderada;
 - Linear por Regressão;
 - Média;
 - Mediana;
 - Não linear;



Sistemas de Múltiplos Preditores

- Pesquisa atual do Doutorado:
 - Motivação:
 - Através de combinação os modelos ruins do Ensemble podem atrapalhar a previsão final;
 - O oracle apresenta desempenho melhor que o single best;
 - Proposta:
 - Criar uma técnica de seleção dinâmica com base no comportamento do oracle;



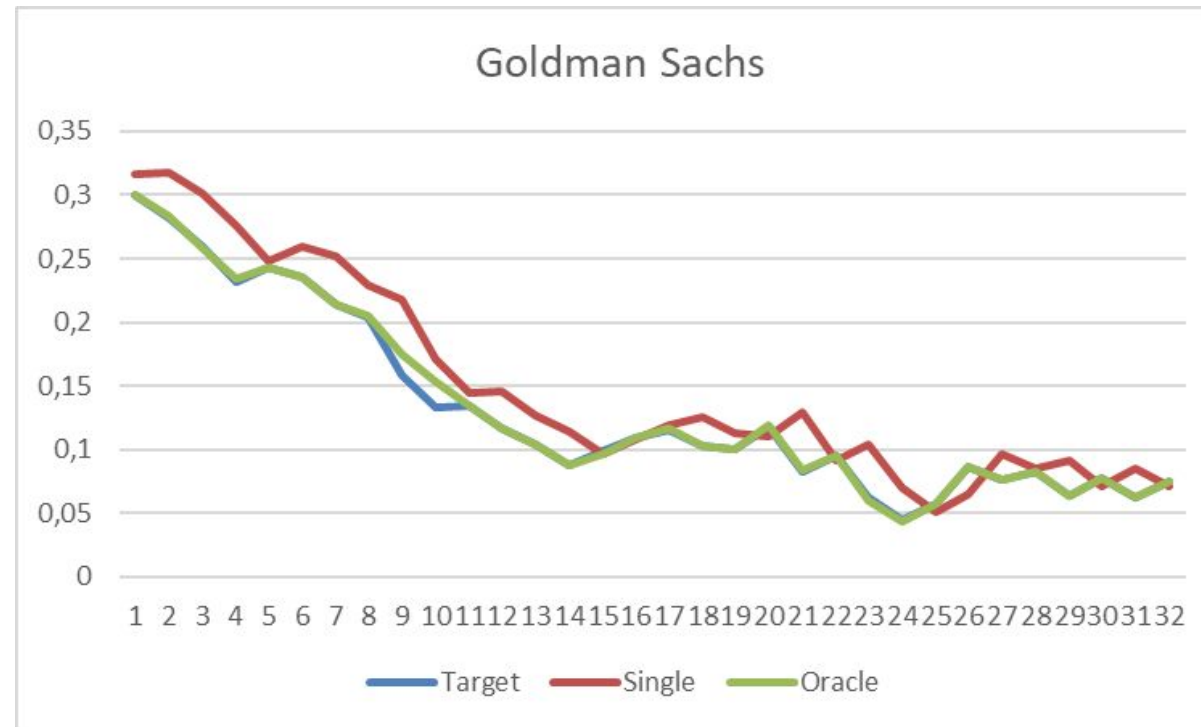
Sistemas de Múltiplos Preditores

- Pesquisa atual do Doutorado:
 - Motivação:
 - Através de combinação os modelos ruins do Ensemble podem atrapalhar a previsão final;
 - O oracle apresenta desempenho melhor que o single best;
 - Proposta:
 - Criar uma técnica de seleção dinâmica com base no comportamento do oracle;



Sistemas de Múltiplos Preditores

- Pesquisa atual do Doutorado:



Machine Learning para previsão de Séries Temporais

Prof. Eraylson Galdino