

Online sequential extreme learning machine with kernels for nonstationary time series prediction

Xinying Wang, Min Han
Artigo da Neurocomputing - 2014



UNIVERSIDADE
FEDERAL
DE PERNAMBUCO



Análise e Previsão de Séries Temporais

Paulo Henrique Muniz Ferreira (*Aluno*)
Adriano Lorena Inácio Oliveira (*Professor*)

- 1 Introdução
- 2 Extreme Learning Machine
- 3 Experimentos
- 4 Conclusão

1 Introdução

2 Extreme Learning Machine

3 Experimentos

4 Conclusão

Previsão de séries temporais:

- tem desempenhado um papel crucial no desenvolvimento de técnicas para a modelagem do sistema dinâmico
- previsão de séries temporais não-lineares

Séries temporais não-lineares:

- é necessário uso de técnicas de processamento de sinais não-linear
- Máquina de vetores de suporte (SVM), Redes Neurais (RN) e outros métodos de aprendizagem de máquina tem sido aplicados para previsão de séries temporais

Séries Temporal Não Estacionária

É quando suas propriedades estatísticas (ex. média, variância ou covariâncias) se modifica a medida que modificamos as variáveis no tempo

- Na prática, as séries são quase sempre não-estacionárias
- Em algumas situações, isso pode restringir os métodos estacionários

1 Introdução

2 Extreme Learning Machine

3 Experimentos

4 Conclusão

Extreme Learning Machine - ELM

É uma técnica de aprendizagem para rede neurais do tipo *feedforward* de uma única camada intermediária baseada na minimização da norma dos pesos de saída

Objetivos da ELM

- Convergência Rápida
- Pouca quantidade de parâmetros a serem ajustados

Extreme Learning Machine com Kernel

- Para evitar o problema da seleção do número de nodos da camada intermediária, o ELM com kernel foi desenvolvido
- É removido o mapeamento da camada intermediária $\mathbf{h}(x)$
- É adicionado o mapeamento por função de kernel $\phi(x)$ na máquina de vetor de suporte

Extreme Learning Machine com Kernel

A Matriz de Kernel da ELM

$$K_{ELM} = HH^T \quad (1)$$

$$K_{ELM(i,j)} = \mathbf{h}(\mathbf{x}_i) \cdot \mathbf{h}(\mathbf{x}_j) = K(\mathbf{x}_i, \mathbf{x}_j) = K_{i,j} \quad (2)$$

A função de saída

$$f(\mathbf{x}) = \mathbf{h}(\mathbf{x})H^T(HH^T + \frac{I}{C})^{-1}\mathbf{t} = \begin{bmatrix} K(\mathbf{x}, \mathbf{x}_1) \\ \vdots \\ K(\mathbf{x}, \mathbf{x}_N) \end{bmatrix}^T (K_{ELM} + \frac{I}{C})^{-1}\mathbf{t} \quad (3)$$

Extreme Learning Machine com Kernel - *Online* e Sequencial

- Ambas ELM e ELMK são métodos de aprendizagem em batch e *offline*
- Podem ter dificuldades computacionais para grandes conjuntos de treinamento
- Em muitas situações reais, é necessária uma aprendizagem em tempo de execução
 - A medida que aparece um novo padrão, a rede deverá aprender incrementalmente
- O artigo apresenta uma proposta de modificação do ELMK para aprendizagem sequencial e *online*

Modo offline

$$R = (K_{ELM} + \frac{I}{C})^{-1} = \left(\begin{bmatrix} K_{1,1} & \cdots & K_{1,N} \\ \vdots & \ddots & \vdots \\ K_{N,1} & \cdots & K_{N,N} \end{bmatrix} + \frac{1}{C} \right)^{-1} \quad (4)$$

Solução online

$$R^{new} = \left(R^* + \frac{1}{C} \right)^{-1} \quad (5)$$

$$R^* = \begin{bmatrix} K_{ELM} & K'_{bs} \\ (K'_{bs})^T & K''_{bs} \end{bmatrix} \quad (6)$$

Modo offline

$$K'_{bs} = \left(\begin{bmatrix} K'_{1,1} & \cdots & K'_{1,bs} \\ \vdots & \ddots & \vdots \\ K'_{N,1} & \cdots & K'_{N,bs} \end{bmatrix} \right)^{-1} \quad (7)$$

onde $K'_{k,i} = K(\mathbf{x}_k, \mathbf{x}_i^{new})$

Solução online

$$K''_{bs} = \left(\begin{bmatrix} K''_{1,1} & \cdots & K''_{1,bs} \\ \vdots & \ddots & \vdots \\ K''_{bs,1} & \cdots & K''_{bs,bs} \end{bmatrix} \right)^{-1} \quad (8)$$

onde $K''_{k,i} = K(\mathbf{x}_k^{new}, \mathbf{x}_i^{new})$

Previsão de Série Temporal Não-Estacionária

- Séries temporais não-estacionária existe em diversos problemas reais
 - sinal de voz
 - séries temporais financeiras
 - séries temporais meteorológicas
- Aprendizagem online pode aprender novas amostras sem a necessidade de retreinar todo conjunto de treinamento
- e tem sido uma ferramenta preferida para a previsão de séries temporais não-estacionárias

Previsão de Série Temporal Não-Estacionária

- Um paradigma em memória fixa é projetado para melhorar a previsão de série temporal não-estacionária
- A ideia é atualizar o modelo de previsão em tempo real para atender a dinâmica mais recentes da série temporal não-estacionária
- Para isso, o autor propõe usar um conjunto de treinamento adaptativo

Previsão de Série Temporal Não-Estacionária

- há um conjunto de treinamento em uma memória fixa com tamanho máximo pré-estabelecido
- um algoritmo de aprendizagem incremental é realizado
- é removido as $|O|$ amostras mais antigas
- é adicionado as bs novas amostras

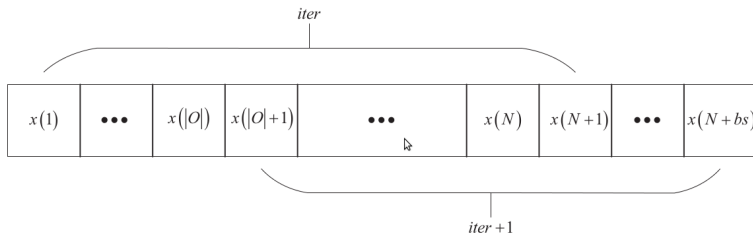


Fig. 1. Fixed memory prediction scheme for nonstationary time series prediction.

Matematicamente

$$R^{new} = R_{I,I} - R_{I,O}(R_{O,O})^{-1}R_{k,O} \quad (9)$$

onde $I = [1 \cdots k (k + |O| + 1) \cdots N]$ e O é o conjunto de amostras removidas

1 Introdução

2 Extreme Learning Machine

3 Experimentos

4 Conclusão

- Foram feitos experimentos com 5 séries temporais
 - a série temporal caótica de Mackey-Glass misturada com uma senoide
 - a série temporal do número anual de manchas solares
 - a série temporal dos índices da *Dow Jones Industrial Average* (DJIA)
 - a série temporal S&P500
 - a série temporal Santa Fe - *A:Laser generated data*

- Foi usada a normalização *z-score*
 - a média fica igual à zero e o desvio padrão, igual a um
- O número de valores prévios usados para previsão foi $m = 1$
- O tamanho da memória fixa foi 100 amostras
- $|O|$ igual ao número de novas amostras ($|O| = bs$).

Configuração dos Experimentos

- É realizado previsão de uma passo a frente
- O OS-ELMK foi avaliado usando o treinamento incremental
 - um-por-um
 - bloco-por-bloco com o bloco de tamanhos 10 e 20
- O ELMK foi realizado utilizando uma aprendizagem *online* aproximada.
- Como memória fixa é 100 amostras
 - as primeiras 100 amostras não são previstas
 - a previsão é avaliado com as amostras restantes

- A fim de demonstra a eficácia foram avaliadas as seguintes medidas
 - root mean squared error (RMSE)
 - normalized root mean square error (NRMSE)
 - mean absolute percentage error (MAPE)
 - symmetric mean absolute percentage error (SMAPE)

Resultados - Série Mackey–Glass Não Estacionária

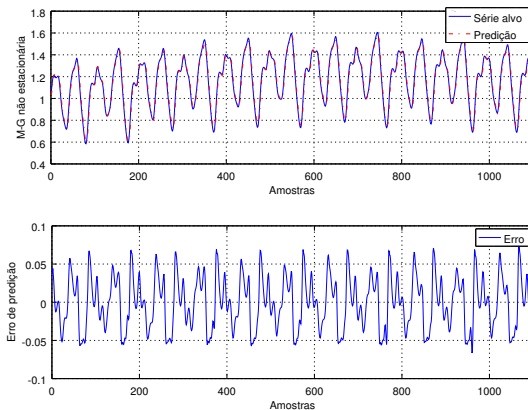


Tabela: Os resultados dos experimentos na previsão da série temporal caótica Mackey–Glass não estacionária.

Método	Batch Size	Tempo de Aprendizagem	RMSE	NRMSE	MAPE	SMAPE
ELMK	-	146.668	0.0336	0.0328	2.563	0.0256
	1	3.989	0.0337	0.0330	2.600	0.0260
OS-ELMK	10	0.477	0.0351	0.0344	2.693	0.0269
	20	0.291	0.0356	0.0349	2.715	0.0271

Resultados - série temporal dos índices da DJIA

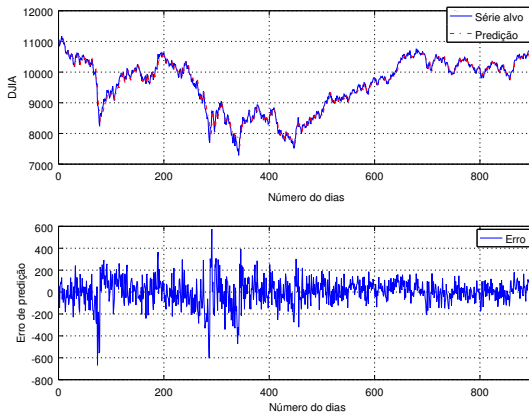


Tabela: Os resultados dos experimentos na previsão da série temporal DJIA.

Método	Batch Size	Tempo de Aprendizagem	RMSE	NRMSE	MAPE	SMAPE
ELMK	-	84.575	118.1758	0.0304	0.9164	0.0091
	1	2.995	122.2815	0.0314	0.9734	0.0097
OS-ELMK	10	0.347	184.5278	0.0474	1.2019	0.0119
	20	0.247	208.1550	0.0535	1.3490	0.0133

Resultados - série temporal S&P500

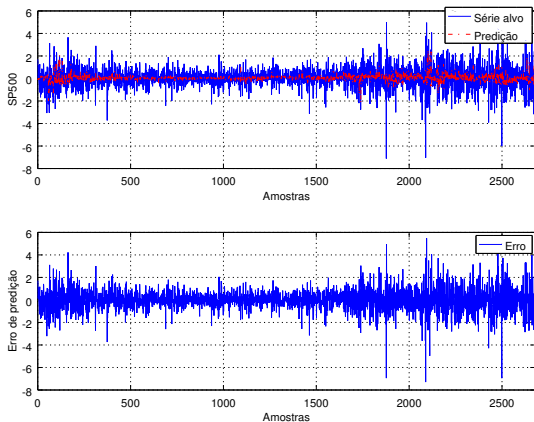


Tabela: Os resultados dos experimentos na previsão da série temporal SP500.

Método	Batch Size	Tempo de Aprendizagem	RMSE	NRMSE	SMAPE
ELMK	-	1628.208	0.9531	0.0788	1.7543
	1	11.115	0.9961	0.0823	1.6071
OS-ELMK	10	1.121	0.9914	0.0819	1.6021
	20	0.604	0.9936	0.0821	1.6031

1 Introdução

2 Extreme Learning Machine

3 Experimentos

4 Conclusão

- Uma versão incremental e sequencial da ELM com kernels (OS-ELMK) foi implementada e estudada
- Foi estudado uma técnica de memória fixa para previsão de séries temporais não estacionárias
- O modelo estudado tem as seguintes vantagens e desvantagem:
 - aprendizado sequencial e incremental
 - adaptação para um algoritmo de aprendizagem decremental
 - desempenho semelhante em termos de erro com um tempo de aprendizagem mais curto em comparação
 - dificuldade em rastrear adequadamente séries com volatilidade

Online sequential extreme learning machine with kernels for nonstationary time series prediction

Xinying Wang, Min Han
Artigo da Neurocomputing - 2014



UNIVERSIDADE
FEDERAL
DE PERNAMBUCO



Análise e Previsão de Séries Temporais

Paulo Henrique Muniz Ferreira (*Aluno*)
Adriano Lorena Inácio Oliveira (*Professor*)