



DMC_x^2 e aprendizado de máquina aplicados à análise de dados de séries temporais do clima

Apresentação do andamento da pesquisa

Fernando Ferraz Ribeiro

15/05/2023

UEFS PPGM

1. Introdução
2. Fundamentação Teórica
3. Resultados
4. Considerações Finais
5. Referências

Introdução

Objetivo Principal

Investigar as correlações entre as variáveis climáticas através do coeficiente DMC_x^2 e utilizar o conhecimento destas correlações para alimentar um modelo preditivo do clima.

Objetivos Gerais

1. Implementar um algoritmo computacional geral para calcular o DMC_x^2 para qualquer número de séries temporais.
2. Analisar um conjunto de dados climáticos contendo medições meteorológicas de todas as capitais brasileiras.
3. Analisar um conjunto de dados meteorológicos sobre radiação solar com estações locadas em diversas partes do globo.
4. Desenvolver e implementar um algoritmo de predição baseado em aprendizado de máquina e redes neurais artificiais agregados com o coeficiente DMC_x^2 .

Fundamentação Teórica

Os dados que tem resolução do município e englobam uma parcela significativa dos municípios da Bahia, serão representados no arquivo dos municípios da Bahia. Conjuntos de dados que apresentam escala dos estados, serão representados no mapa dos estados do Brasil, Conjunto que apresentem apenas as capitais e cidades maiores, representaremos nos mapas dos municípios brasileiros.

DFA - Algoritmo

1. Pegando a série temporal $\{x_i\}$ com i variando de 1 à N , a série integrada X_k é calculada por $X_k = \sum_{i=1}^k [x_i - \langle x \rangle]$ com k também variando entre 1 e N ;
2. A série X_k é dividida em $N - n$ caixas de tamanho n (escala temporal), cada caixa contendo $n + 1$ observações, iniciando em i até $i + n$;
3. Para cada caixa um polinômio (geralmente de grau 1) é ajustado, gerando $\tilde{X}_{k,i}$ with $i \leq k \leq (i + n)$ eliminando assim a tendência (detrended values);
4. para cada caixa é calculado: $f_{DFA}^2(n, i) = \frac{1}{1+n} \sum_{k=i}^{i+n} (X_k - \tilde{X}_{k,i})^2$
5. Para todas as caixas de uma escala temporal o DFA é calculado como: $F_{DFA}(n) = \sqrt{\frac{1}{N-n} \sum_{i=1}^{N-n} f_{DFA}^2(n, i)}$;
6. Para um número de diferentes escalas temporais (n), com valores possíveis entre $4 \leq n \leq \frac{N}{4}$, a função F_{DFA} é calculada para encontrar a relação entre $F_{DFA} \times n$

Resultados

Resultados Esperados

Para cada um dos conjunto de dados que apresentem escala de município e incluam um grande conjunto de cidades do município da Bahia, serão representados em conjuntos de 7 mapas, representando em formato coroplético a mediana da mobilidade de cada município. A necessidade de normalização/padronização dos dados será avaliada caso a caso.

Os mapas que apresentem poucos cidades no território nacional, ou que apresente dados no nível d bairro será utilizada visualização por tamanho de ponto.

Considerações Finais

As mudanças provocadas pela pandemia de COVID 19 em diversas áreas ainda precisa ser melhor entendido. No que diz respeito aos dados oriundos de dispositivos móveis de comunicação muito temos que analisar.

Tanto das possibilidades e limitações, quanto as questões éticas e respeito à privacidade devem ser analisadas. A visualização de dados proposta neste estudo pretende fornecer subsídios para este debate.

Armazenamento das informações coletadas em um banco de dados georeferenciados.

Referências
