

Avaliação final - Aprendizado Dinâmico

Para esta avaliação, você deve entregar três arquivos:

- um arquivo de dados em .csv com a base resultante das análises
- um relatório em .ipynb com suas análises comentadas
- um arquivo .pdf gerado a partir do .ipynb

Os dados de casos confirmados e mortes diárias por COVID-19 nos Estados Unidos estão disponíveis no arquivo covid19USA.csv.

1. Você deve fazer a leitura dos dados, estabelecer um índice por datas, verificar se os dados estão corretos e realizar algum procedimento necessário caso haja dados faltantes.
2. Proponha uma decomposição dos dados de mortes em tendência e sazonalidade. Obtenha os gráficos de autocorrelação e autocorrelação parcial. Comente brevemente os resultados.
3. Divida a base em treino e teste, com 21 observações na base de teste. Você deverá apresentar em distintas colunas na mesma base de dados que será entregue:
 - Os dados originais de casos confirmados e mortes.
 - As previsões de Holt e Holt Winters para a variável mortes.
 - As previsões obtidas pelo melhor modelo SARIMA, com ordens selecionadas segundo critério preestabelecido para a variável mortes. Identifique as componentes desse melhor modelo (por exemplo, “o modelo tem uma componente autorregressiva não sazonal de ordem .., uma componente de médias móveis não-sazonal de ordem..., etc).
 - As previsões obtidas por um modelo de redes dinâmicas para a variável mortes.
4. Utilize o erro quadrático médio e erro absoluto médio para comparar as previsões obtidas com as observações da base de teste.
5. Considere o método que produziu a melhor previsão para os dados de teste e faça a previsão para mais 14 dias (além das observações da base de dados). Essa previsão pode ser apresentada com um gráfico nos arquivos .ipynb e .pdf.