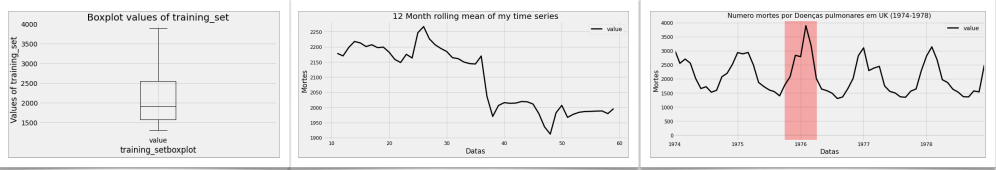
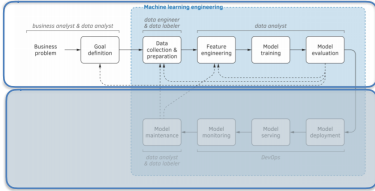
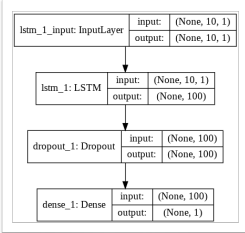
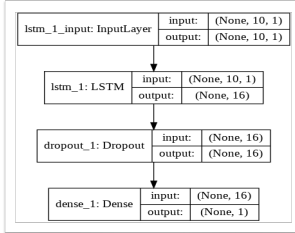
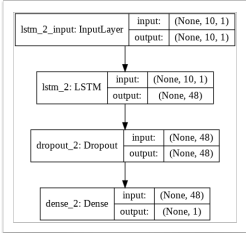
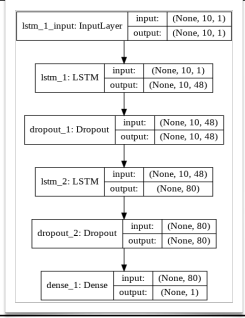
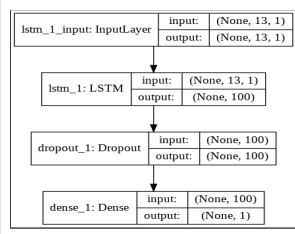
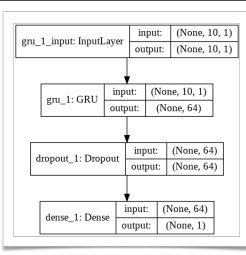
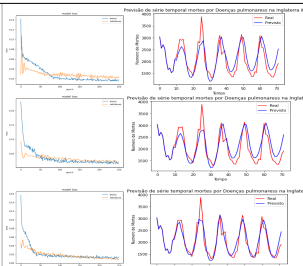
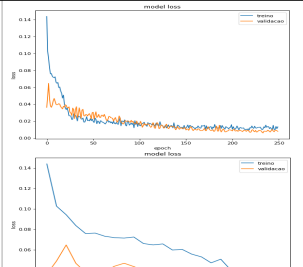
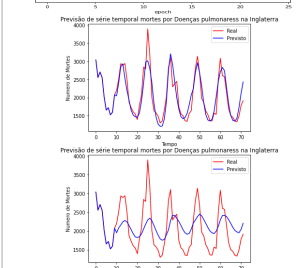
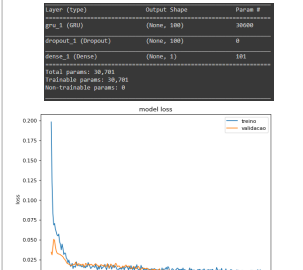
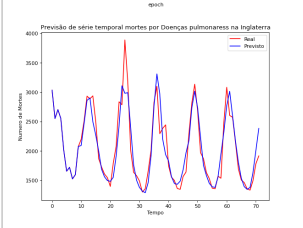




Trabalho Final RN_Previsão de Séries e Inferência por redes neurais /Estudo de Caso				
Nome do Projeto:	Projeto de Conclusão da disciplina RN Previsão de Séries e Inferência por redes neurais			
Aprovação:	Professor PhD Leonardo Mendonça Forero	Revisão	Data	
Autor	Rodolfo Stocker			
Aplicação	Previsão de Séries temporais da disciplina RN 2019-4 BI			
Antecedentes do Projeto: EDA	<p>Dataset : y:#(inter) Ideaths (variável discreta) , é uma série temporal das mortes mensais por doenças pulmonares no Reino Unido entre 1974-1978. Problema do tipo supervisionado com dados sequenciais. Input sequência de números. many to one (univariável: ou uma variável de resposta).</p> <p>Nos gráficos, com as visualizações da série temporal pode-se obter alguns insights como:</p> <ul style="list-style-type: none"> - A maior e menor mortalidade mensal foi registrada em 1976 (total incluindo ambos os sexos). - Tem havido uma flutuação constante nos números ao longo dos meses (em vários anos), mas o comportamento geral tem sido consistente (Sazonalidade). - A tendência para o final de 1978 mostra no gráfico Média móvel que os números estão diminuindo. <p>Decomposição da série temporal:</p> <p>Ao visualizar dados de série temporal, confira alguns padrões distinguíveis:</p> <ul style="list-style-type: none"> - Sazonalidade: os dados mostram um padrão periódico claro? É verificado a partir do próprio gráfico. - Tendência: os dados seguem uma inclinação consistente para cima ou para baixo? É verificado, em seguida, as médias móveis são calculadas - Ruído: existem pontos discrepantes ou valores ausentes que não são consistentes com o resto dos dados? sim, no boxplot, a presença de outliers é confirmada. <p>**Poderia aprimorar ainda mais as métricas, logo de avaliar os modelos a partir dos dados de Sazonalidade isolados de ruído e tendência, mas quando devo manter uma anomalia (outliers)?**</p> <p>Ivneria:</p> <pre>from statsmodels.tsa.seasonal import seasonal_decompose # decomposition = sm.tsa.seasonal_decompose(time_series)</pre>			
EDA				
Scopo do fluxo de Trabalho da disciplina RN:	<p>Workflow Pipeline:</p> <ul style="list-style-type: none"> - Goal Definition : Metodologia** - Data Preparation & Featuring Eng: Pré-processamento (Normalização-Janelamento- Reshape) dataset treino-teste - Construção da Topologia da rede neural LSTM (fcamadas(neurônios)) + FullDense (1). - Compilação do modelo (optimizer='Adam', 'SGD'). - Model Training: Treinamento do Modelo .fit(). - Model Evaluation: Avaliação das Métricas de regressão e visualização das curvas Loss-Epoch, previsão das series (y_Treino ,y_Test). 			
Objetivo dos estágios do Trabalho da disciplina RN:	<p>Método busca Iterativa Manual***.</p> <p>Busca iterativa por uma solução ideal ou RMSE mais baixo em um espaço de busca (Ψ) guiado por a intuição e que fornece algum grau de compreensão sobre Ψ(espaço de busca). Selecionei hiperparâmetros com base na intuição , experimentação , suposição, treinei o modelo com os hiperparâmetros e pontuei nos dados de validação, usei CallBack, mc (modelcheckpoint) e repeti o processo até ficar satisfeito com os resultados.</p> <p>*** Existem outras três abordagens para ajustar os hiperparâmetros de um modelo de aprendizado da máquina que não sejam parte do scope deste trabalho.</p> <p>Grid Search: Exaustivo ou Abrangente , nesta abordagem, todas as combinações de valores de hiperparâmetros são testadas.</p> <p>configura uma grade de valores de hiperparâmetros e, para cada combinação, treine um modelo e pontue nos dados de validação.</p> <p>Random Search : O número de iterações de pesquisa é definido com base no tempo / recursos. Configurei uma grade de valores de hiperparâmetros e selecione combinações aleatórias para treinar o modelo e pontuar.</p> <p>Automated Hyperparameter Tuning: Métodos como gradiente descendente, otimização bayesiana ou algoritmos evolutivos para realizar uma busca guiada pelos melhores hiperparâmetros.</p>			
Entregáveis:	<p>Documentação do projeto:</p> <ul style="list-style-type: none"> - Relatório no formato tipo word ou pdf. - Colab Screenshot. - Solução proposta. S** no arquivo anexo se precisar Notebooks Colab RStocker_BI2019_4_RNN_IDisciplina_IterGRU.ipynb. 			
Alcance do Projeto:	<p>A abordagem do trabalho iniciará orientando a preparação do dataframe, passando pelo tratamento dos dados (normalização-escalamento-janelamento-reshaping). A interação das variantes topológicas das redes neurais propostas seguida da Inferência e pós-processamento (loop de avaliações iterativas), ou seja ,</p> <p>Otimizar as métricas : RMSE (Min) - MAPE (Min) - R2 - (Max < 1)</p> <p>Diagnosticando o comportamento do modelo por meio das curvas de aprendizado Loss-Epoch: Underfit, Overfit, Good fit.</p> <p>Curva de aprendizado do treino: curva de aprendizado foi calculada a partir do conjunto de dados de treinamento que deram uma ideia de como o modelo está aprendendo.</p> <p>Curva de aprendizado de validação: curva de aprendizado calculada a partir de um conjunto de dados de validação que dá uma ideia do quão bem o modelo está generalizando.</p> <p>Curvas de ajuste da variável resposta predici da serie temporal.</p> <p>A abordagem final do trabalho se orientou pela busca proposital da solução ótima aceitável, mesmo dentro, como fora do scope do modelo proposto que forneceu uma solução mais aprimorada que as soluções anteriormente avaliadas.</p>			
Informação disponibilizada:	<p>Trabalho sugerido e dataset proposto da disciplina.</p> <ul style="list-style-type: none"> - Notebooks Colab. - Base de Treino no arquivo : Treino.csv. - Base de Teste no arquivo: Test.csv. 			

Trabalho Final RN: Previsão de Séries e Inferência por redes neurais /Estudo de Caso					
Nome do Projeto:	Projeto de Conclusão da disciplina RN Previsão de Séries e Inferência por redes neurais				
Aprovação:	Professor PhD Leonardo Mendonça Forero	Revisão	Data		
Autor	Rodolfo Stocker				
Aplicação	Previsão de Séries temporais da disciplina RN 2019-4 B1				
<p>Construção do Modelo Original RN LSTM com Otimizador SGD, uma camada (100) neurônios + FullDense (1).</p> <p>Construção do Modelo Original RN LSTM com Otimizador Adam, uma camada (16) neurônios + FullDense (1).</p> <p>Construção do Modelo Original RN LSTM com Otimizador Adam, uma camada (48) neurônios + FullDense (1).</p>	  				
	  				
	<p>Construção do Modelo Original RN LSTM com Otimizador Adam, uma camada (48) + Uma Camada oculta (80) neurônios + FullDense (1).</p> <p>Construção do Modelo Original RN LSTM com Janelamento 13, Otimizador Adam, uma camada (100) neurônios + FullDense (1).</p> <p>Construção do Modelo RN, GRU com Otimizador Adam, uma camada (64) neurônios + FullDense (1).</p>				
Critério de êxito:	Enviei toda a documentação do projeto dentro do período pré - definido de 04 de Abril de 2021, disponibilizando as informações para revisão.				
***Metodologia:	<p>Referencia os trechos do texto original definição do problema " A base apresentada é de pessoas mortas por doenças pulmonares na Inglaterra, a ideia é prever em um período de tempo com redes LSTM. Por favor rodem o código com a base de dados, vejam os resultados e façam as seguintes mudanças".</p> <p>Iterações:</p> <ol style="list-style-type: none"> 1. "Mude o otimizador de 'adam' para 'sgd' dentro da configuração da rede e veja os resultados de RMSE". 2. "Voltando a configuração inicial mude o número de neurônios da camada LSTM, primeiro para 16 e depois para 48 e veja o resultado do RMSE." 3. "Coloque uma segunda camada LSTM e veja o resultado do RMSE". 4. "Mude por dois valores que você ache conveniente as épocas de treinamento e veja o resultado do RMSE". 5. "Com as mudanças feitas proponha sua configuração ideal e Apresente um relatório com os resultados em word e explique como chegou a sua configuração ideal". 6. "Crie e rode o código da rede GRU e repare os resultados de RMSE, MAPE, R2" com um janelamento [50,10,1], otimize com Adam que forneceu as melhores métricas".** 				
Conclusão :	<p>O problema foi direcionado para a busca manual e seletiva baseada em experimentação de variantes tipológicas do modelo LSTM (Vanilla,Stacked,GRU) e na avaliação dos resultados fornecidos pelos modelos FullDense preditivos e minimizando erros das previsões.</p> <p>Sabemos que os modelos estatísticos funcionam para a previsão de séries temporais. No entanto, existem algumas limitações para esses métodos:</p> <ul style="list-style-type: none"> - Precisam de dados completos para o treinamento - Normalmente, lidam com conjunto de dados univariados <p>Os métodos de aprendizagem profunda como no escopo do problema proposto, são capazes de lidar com os desafios acima:</p> <ul style="list-style-type: none"> -Facilidade de incorporação de variáveis exógenas (aplicar em conjuntos de dados univariados e multivariados) -Além do anterior, captura interações de atributos não lineares. <p>O modelo LSTM original Vanilla é composto de uma única camada LSTM, seguida por uma camada de saída feedforward padrão. O modelo Stacked LSTM é uma extensão deste modelo que possui várias camadas LSTM ocultas, onde cada camada contém várias células de "memória ou estados celulares". Portanto deveriam conseguir melhores métricas de ajuste e robustez para lidar com problemas, onde existe o contexto de longo prazo é sempre o que se disponha de suficientes dados para treinar e teste.</p> <p>Ref. Uma rede neural profunda [DNN] é uma rede neural artificial [ANN] com várias camadas entre as camadas de entrada e saída.</p>				

Trabalho Final RN, Previsão de Séries e Inferência por redes neurais / Estudo de Caso					
Nome do Projeto:	Projeto de Conclusão da disciplina RN Previsão de Séries e Inferência por redes neurais				
Aprovação:	Professor PhD Leonardo Mendonça Fereira				
Autor:	Rodolfo Stocker				
Aplicação:	Previsão de Séries temporais da disciplina RN 2019-4 BI				
Avaliação dos Modelos Original RN LSTM com Otimizador SGD, uma camada (100) neurônios + FullDense (1).	Diagnóstico da curva de aprendizado Iterações 1 e 2: Além das métricas de ajustes RMSE, e uma boa prática, até ser recomendável a visualização das curvas de aprendizado, por meio, da convergência ou divergência das curvas de treino e avaliação através dos ajustes dos hiperparâmetros feitos em cada interação. Um conjunto de dados de validação não representativo significa que o conjunto de dados de validação não fornece informações suficientes para avaliar a capacidade do modelo para generalizar. Isto pode ocorrer se o conjunto de dados de validação tiver poucos exemplos em comparação com o conjunto de dados de treinamento. Também pode ser identificado por uma perda de validação menor do que a perda de treinamento. Nesse caso, indica que o conjunto de dados de validação pode ser mais fácil para o modelo prever do que o conjunto de dados de treinamento.				
	Avaliação do Modelo Original RN LSTM com Otimizador Adam, uma camada (16) neurônios + FullDense (1).	Avaliação das métricas e interpretação de RMSE (Min) R2 (Max < 1) Root Mean Square Error (RMSE) é o desvio padrão dos resíduos (erros de predição). ele informa o quão concentrados os dados estão em torno da linha de melhor ajuste: "Quem obtiver o menor RMSE possui a curva mais próxima dos pontos". O RMSE para o modelo RN LSTM (SGD) foi 377 mortes, MAPE 21%, R2 32% o modelo explica 31%.			
	Avaliação do Modelo Original RN LSTM com Otimizador Adam, uma camada (48) neurônios + FullDense (1).	O RMSE para o modelo RN LSTM (16 Neurônios) foi 418 mortes, MAPE 22%, R2 46% o modelo explica 27%. O RMSE para o modelo RN LSTM (48 Neurônios) foi 256 mortes, MAPE 11%, R2 78% o modelo explica 78%. R2: A porcentagem de variação na resposta que é explicada pelo modelo, para o número de preditores do modelo em relação ao número de observações. A partir da progressão dos gráficos de Previsão, embora falte alguns picos e depressões, é capaz de ir capturando as sazonalidades com os ajustes do pré-processamento e hiperparâmetros.			
Avaliação do Modelo Original RN LSTM com Otimizador Adam, uma camada (48) + Uma Camada oculta (80) neurônios + FullDense (1) e Treino Epoch (250, 25)	Iterações 3 e 4 → Mantive de otimizador "Adam" da configuração original, o número de neurônios na primeira camada 48 e na segunda camada de 80 neurônios com duas iterações com ciclos de treino 250 e 25 épocas respectivamente. No gráfico acima ou no screenshot da lado, mostra uma estrutura do RNN LSTM com duas camadas. A seguir, vamos ver no detalhe o modelo a implementar. Além da camada RN LSTM com 48 neurônios, aplicamos uma segunda camada RN LSTM stacked ou seja a camada oculta de 80 neurônios, e uma camada FullDense para previsão de dados de uma série temporal univariável (many to one). O modelo LSTM original Stacked é composto de uma única camada LSTM oculta, seguida por uma camada de saída feedforward padrão, onde a camada contém várias células de "memória ou estados celulares". Portanto, deveriam conseguir melhores métricas de ajuste e robustez para lidar com problemas, onde existirem o contexto de longo prazo é sempre o que se dispõem de suficientes dados para treino e teste. A partir da topologia da rede neural é as iterações 3-4 e da compilação do modelo LSTM Stacked, o Modelo de ajuste foi treinado (fit), visualizando convergência das curvas de aprendizagem treino/teste (Loss-Epoch) para visualizar a presença de underfitting ou overfitting nos gráficos das curvas de ajuste: treino/teste de previsão e interprete seu significado tanto com 250 épocas de treino e 25 épocas de teste. Diagnóstico da curva de aprendizado Iterações 3 e 4: Um conjunto de dados de validação não representativos significa que o conjunto de dados de validação não fornece informações suficientes para avaliar a capacidade do modelo para generalizar. Isto pode ocorrer se o conjunto de dados de validação tiver poucos exemplos em comparação com o conjunto de dados de treinamento. Também pode ser identificado por uma perda de validação menor do que a perda de treinamento. Nesse caso, indica que o conjunto de dados de validação pode ser mais fácil para o modelo prever do que o conjunto de dados de treinamento. Este caso pode ser identificado por uma curva de aprendizado para perda de treinamento que parece um bom ajuste e uma curva de aprendizado para perda de validação que mostra movimentos ruidosos em torno da perda de treinamento.				
	Este caso pode ser identificado por uma curva de aprendizado para perda de treinamento que parece um bom ajuste e uma curva de aprendizado para perda de validação que mostra movimentos ruidosos em torno da perda de treinamento. Avaliação das métricas e interpretação de RMSE (Min), R2 (Max < 1) Root Mean Square Error (RMSE) é o desvio padrão dos resíduos (erros de predição). ele informa o quão concentrados os dados estão em torno da linha de melhor ajuste: "Quem obtiver o menor RMSE possui a curva mais próxima dos pontos". Iter 3. O RMSE do RN LSTM de duas camadas (48, 80 neurônios): 241 mortes, MAPE 9,6%, R2 80% o modelo explica 80%. Iter 4. O RMSE do RN LSTM de duas camadas (48, 80 neurônios): 518 mortes, MAPE 27%, R2 8,9%, o valor significa que o modelo explica uma % negativa da variável de resultado, o que não é uma sugestão intuitivamente razoável.				
	R2: A porcentagem de variação na resposta que é explicada pelo modelo, para o número de preditores do modelo em relação ao número de observações. A partir do gráfico de Previsão Iteração 3, embora falte alguns picos e depressões, é capaz de capturar as sazonalidades com os ajustes do pré-processamento e hiperparâmetros.				
Construção e avaliação do Modelo RN GRU com Otimizador Adam, uma camada (44) neurônios + FullDense (1)	Alé agora, além dos modelos RN LSTM original Vanilla, composto de uma única camada LSTM oculta, seguida por uma camada de saída feedforward padrão. O modelo Stacked LSTM é uma extensão deste modelo que possui várias camadas LSTM ocultas, onde cada camada contém várias células de "memória ou estados celulares". Portanto, deveriam conseguir melhores métricas de ajuste e robustez para lidar com problemas, pois, onde existe o contexto de longo prazo é sempre o que se dispõe de suficientes dados para treino e teste. Conclusão interação GRU: A RNN GRU, com uma camada de 64 neurônios, seguida por uma camada de saída feedforward padrão (1 neurônio) performou melhores métricas RMSE, MAPE, R2 em comparação aos outros tradeoff avaliados LSTM (Vanilla e Stacked) e com a capacidade de generalização para dados não vistos pela rede e com um custo computacional menor (menos parâmetros para treinar). No gráfico, o conjunto de dados de validação inicia embora do treino o que indicaria que não é representativo, logo, as épocas 25. As curvas de aprendizado de treinamento e validação mostram um conjunto de dados de validação que se ajustam bem e conseguem generalizar. O RMSE para o modelo RN LSTM (SGD) foi 219 mortes, MAPE 7,6%, R2 84% o modelo explica 84%. Resíduos são uma medida do quão longe os pontos de dados da linha de regressão estão. O RMSE é uma medida do quão espalhados são esses resíduos em torno da linha de melhor ajuste. Minimizando o erro associado à curva ajustada. Quem obtiver o menor RMSE possui a curva mais próxima dos pontos. Métrica Percentual Absoluta do Erro (MAPE - Mean Absolute Percentage Error) Expressa a acurácia do erro em porcentagem. No caso temos um MAPE de 7,6%, quer dizer que o nosso ajuste está errado em 7,6% dos dados. R2 é a porcentagem de variação na resposta que é explicada pelo modelo 84%, para o número de preditores (1) do modelo em relação ao número de observações. E uma boa prática até ser recomendável, além das métricas de ajustes RMSE, a visualização das curvas de aprendizado, por meio, das mudanças de performance ou convergência das curvas de treino e avaliação, através dos ajustes dos hiperparâmetros feitos em cada interação. O problema de detecção de anomalias para séries temporais é geralmente formulado como a descoberta de pontos de dados discrepantes em relação a algum nível padrão ou usual. O análise o desvio do resíduo e introduz algum limite para ele, obtendo um algoritmo de detecção de anomalias. Nesse caso particular o objetivo foi avaliar o grau do conhecimento conceitual do RN LSTM Forecasting da disciplina.				
	A RNN GRU, com uma camada de 64 neurônios, seguida por uma camada de saída feedforward padrão (1 neurônio) performou melhores métricas RMSE, MAPE, R2 em comparação aos outros tradeoff avaliados LSTM (Vanilla e Stacked) e com a capacidade de generalização para dados não vistos pela rede e com um custo computacional menor (menos parâmetros para treinar).				
	No gráfico, o conjunto de dados de validação inicia embora do treino o que indicaria que não é representativo, logo, as épocas 25. As curvas de aprendizado de treinamento e validação mostram um conjunto de dados de validação que se ajustam bem e conseguem generalizar. O RMSE para o modelo RN LSTM (SGD) foi 219 mortes, MAPE 7,6%, R2 84% o modelo explica 84%. Resíduos são uma medida do quão longe os pontos de dados da linha de regressão estão. O RMSE é uma medida do quão espalhados são esses resíduos em torno da linha de melhor ajuste. Minimizando o erro associado à curva ajustada. Quem obtiver o menor RMSE possui a curva mais próxima dos pontos. Métrica Percentual Absoluta do Erro (MAPE - Mean Absolute Percentage Error) Expressa a acurácia do erro em porcentagem. No caso temos um MAPE de 7,6%, quer dizer que o nosso ajuste está errado em 7,6% dos dados. R2 é a porcentagem de variação na resposta que é explicada pelo modelo 84%, para o número de preditores (1) do modelo em relação ao número de observações. E uma boa prática até ser recomendável, além das métricas de ajustes RMSE, a visualização das curvas de aprendizado, por meio, das mudanças de performance ou convergência das curvas de treino e avaliação, através dos ajustes dos hiperparâmetros feitos em cada interação. O problema de detecção de anomalias para séries temporais é geralmente formulado como a descoberta de pontos de dados discrepantes em relação a algum nível padrão ou usual. O análise o desvio do resíduo e introduz algum limite para ele, obtendo um algoritmo de detecção de anomalias. Nesse caso particular o objetivo foi avaliar o grau do conhecimento conceitual do RN LSTM Forecasting da disciplina.				