



Modelo de LSTM Aplicado na Previsão das Séries de Radiação Solar de Burkina Fasso

Letícia Florentino Pires

JUIZ DE FORA
NOVEMBRO, 2019

Modelo de LSTM Aplicado na Previsão das Séries de Radiação Solar de Burkina Fasso

LETÍCIA FLORENTINO PIRES

Universidade Federal de Juiz de Fora

Instituto de Ciências Exatas

Departamento de Ciência da Computação

Bacharelado em Ciência da Computação

Orientador: Luciana Conceição Dias Campos

JUIZ DE FORA

NOVEMBRO, 2019

MODELO DE LSTM APLICADO NA PREVISÃO DAS SÉRIES DE RADIAÇÃO SOLAR DE BURKINA FASSO

Letícia Florentino Pires

MONOGRAFIA SUBMETIDA AO CORPO DOCENTE DO INSTITUTO DE CIÊNCIAS EXATAS DA UNIVERSIDADE FEDERAL DE JUIZ DE FORA , COMO PARTE INTEGRANTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE BACHAREL EM CIÊNCIA DA COMPUTAÇÃO.

Aprovada por:

Luciana Conceição Dias Campos
Doutor

Leonardo Goliatt da Fonseca
Doutor

Heder Soares Bernardino
Doutor

Gisele Goulart Tavares da Silva
Doutoranda

JUIZ DE FORA
26 DE NOVEMBRO, 2019

A minha mãe e aos meus irmãos.

Ao meu companheiro de vida.

Resumo

O uso de energia fotovoltaica constitui uma alternativa limpa e eficiente ao uso de fontes poluentes de geração de energia elétrica, além de o uso dessa fonte de energia poder chegar a regiões com baixa infraestrutura. No entanto, os custos com o planejamento da instalação são altos, principalmente devido ao uso de sensores de radiação solar, utilizados no estudo de viabilidade do projeto. Para tratar esse problema em Burkina Fasso, o trabalho de Tao et al. [2019] explorou a alternativa de previsão de radiação solar com base nas variáveis ambientais medidas por outros sensores menos custosos, como velocidade do vento e umidade, através do uso de um modelo que não captura as relações temporais da série. Para abordar essa limitação e melhorar os resultados de previsão de radiação solar, o presente trabalho propõe o uso de um modelo *Long Short-Term Memory* LSTM com janelamento dias passados da série temporal. Com essa abordagem chegamos a valores de RMSE entre 0,11 à 0,23 e MAE entre 0,01 à 0,05.

Palavras-chave: Deep Learning, previsão de séries temporais, energia fotovoltaica.

Abstract

The use of photovoltaic energy is a clean and efficient alternative to the use of polluting sources of electricity generation, and the use of this energy source can reach regions with low infrastructure. However, installation planning costs are high, mainly due to the use of solar radiation sensors used in the project feasibility study. To address this problem in Burkina Faso, Tao et al. [2019]'s work explored the solar radiation prediction alternative based on environmental variables measured by other less costly sensors, such as wind speed and humidity, using a model that does not capture the time relationships of the series. To address this limitation and improve the solar radiation prediction results, the present work proposes the use of a LSTM long-windowed *Long Short-Term Memory* model. With this approach we reached RMSE values between 0.11 to 0.23 and MAE between 0.01 to 0.05.

Keywords: Deep Learning, time series forecasting, photovoltaic energy.

Agradecimentos

A minha mãe pela dedicação e força que dedica aos meus irmãos e a mim. Ao meu irmão Marcos pelo constante exemplo. Ao meu irmão Paulo pelos ensinamentos ao longo da vida. Sem vocês, nada disso seria possível.

Ao meu companheiro de vida e profissão, eterno confidente, ser que eu tive a honra de encontrar na minha existência, Gabriel.

Aos professores do Departamento de Ciência da Computação pelos seus ensinamentos em vários aspectos.

Ao professor Andrew Ng, pela sua grande iniciativa educacional, capaz de romper barreiras sócio-educacionais, que culminou nesse trabalho.

"There ain't no such thing as a free lunch."

Robert Heinlein

Conteúdo

Lista de Figuras	7
Lista de Tabelas	8
Lista de Abreviações	9
1 Introdução	10
1.1 Trabalhos relacionados	12
1.2 Organização do trabalho	13
2 Fundamentação Teórica	14
2.1 Série Temporal	14
2.2 Redes Neurais	16
2.2.1 Recurrent Neural Networks	17
2.2.2 Long Short-Term Memory (LSTM)	18
2.3 Redes Neurais de Aprendizagem Profunda (Deep Learning)	21
2.4 Métricas	22
3 Metodologia	24
3.1 Base de dados	24
3.1.1 Pré-processamento dos dados	25
3.2 Modelo LSTM aplicado na previsão das séries de radiação solar de Burkina Fasso	26
4 Estudo de Caso	28
4.1 Configurações de Ambiente	28
4.2 Análise dos experimentos	28
5 Conclusão	33

Lista de Figuras

2.1	Elementos de sazonalidade e tendência da série temporal.	15
2.2	Modelo de neurônio biológico e modelo de neurônio proposto em [Haykin et al., 2009].	16
2.3	Rede neural recorrente. Adaptado de [Cavaioni, 2018].	18
2.4	As células do LSTM em sequência, adaptado de [Colah, 2018]	20
2.5	Célula do LSTM em detalhe, adaptado de [Colah, 2018]	20
2.6	Modelos ANN e Modelo de <i>Deep Learning</i> , adaptado de [de Engenharia, 2018].	21
3.1	Estações de Boromo, Dori, Gaoua e Po, adaptado de [Tao et al., 2019] . . .	24
3.2	Modelo proposto nesse trabalho.	26
4.1	Resultado do conjunto teste Boromo.	31
4.2	Resultado do conjunto teste Dori.	31
4.3	Resultado do conjunto teste Po.	32
4.4	Resultado do conjunto teste Gaoua.	32

Lista de Tabelas

3.1	Variáveis ambientais presentes na base.	25
4.1	Resultados encontrados no conjunto de teste do presente trabalho vs resultados apresentados por [Tao et al., 2019].	29
4.2	Resultados encontrados no conjunto de treino do presente trabalho vs resultados apresentados por [Tao et al., 2019].	31

Lista de Abreviações

DCC	Departamento de Ciência da Computação
UFJF	Universidade Federal de Juiz de Fora
ANN	Artificial Neural Network
DL	Deep Learning
LSTM	Long Short Term Memory
ELM	Extreme Learning Machine
SaD-ELM	Self-Adaptive Evolutionary Extreme Learning Machine
ARIMA	Auto-Regressive-Integrated-Moving-Average models
MAE	Mean Absolut Error
RMSE	Root Mean Squared Error
MAPE	Mean Absolute Percentage Error
VAF	Variance Accounted Factor
SI	Scatter Index
ReLU	Rectified Linear Unit
WS	Velocidade do vento
Tmax	Temperatura máxima
Tmin	Temperatura mínima
Hmax	Umidade máxima
Hmin	Umidade mínima
VPD	Déficit de pressão de vapor
Eo	Evaporação
PV	Photovoltaic
ANEEL	Agência Nacional de Energia Elétrica

1 Introdução

A energia fotovoltaica é obtida por dispositivos que geram eletricidade diretamente da luz solar através de um processo eletrônico que ocorre naturalmente em certos tipos de material, chamados semicondutores. Os elétrons nesses materiais são liberados pela energia solar e podem ser induzidos a trafegar através de um circuito elétrico, alimentando dispositivos elétricos ou enviando eletricidade para a rede [Industries Association, 2019].

Ela é conhecida por ser uma fonte de energia limpa, segura, sustentável e renovável, e é uma opção amplamente usada para substituição de fontes energéticas vindas de combustíveis fósseis [Zhou et al., 2019]. De acordo com o *Global Future Report 2013 REN21*¹ a geração de energia fotovoltaica chegará a 8000 GW² no ano de 2050 [Wang et al., 2019].

A produção de energia solar é intrinsecamente ligada a fatores como radiação solar, temperatura, luminosidade, umidade do ar, velocidade do vento e índices de evaporação. Sendo o mais importante deles a radiação solar, podendo ser usada como variável de predição da produção dos sistemas de energia solar [Arshi et al., 2019].

Os equipamentos de medição de radiação solar apresentam custo elevado e requerem mão de obra especializada para manuseio e manutenção, o que inviabiliza o estudo e instalação de painéis em lugares com recursos financeiros limitados ou mesmo distantes de grandes centros [Tao et al., 2019]. Nesse sentido, os fatores ambientais citados acima podem, em si, ser usados para prever a radiação solar e viabilizar o estudo para instalação de painéis solares, levando energia a regiões onde, geralmente, há um pobre abastecimento de energia elétrica [Tao et al., 2019, Arshi et al., 2019].

Um exemplo de região em que há o uso extensivo de fontes de energia poluentes e pouco eficientes é Burkina Fasso, país localizado na África Subsaariana, onde cerca de 70% da energia gerada vem de fontes de combustíveis termo-fósseis, enquanto que os outros 30% vêm de energia hidroelétrica. Contudo, além de serem fontes poluentes, elas

¹A REN21 é uma comunidade global de energia renovável, cujos membros vêm da academia, dos governos, de ONGs e indústrias com o objetivo de incentivar a transição para as energias renováveis.

²O gigawatt, cujo símbolo é GW, consiste de uma unidade da grandeza física potência.

são susceptíveis a instabilidade do preço do petróleo e têm alto custo, dado a crescente demanda de eletricidade. Além disso, regiões remotas sofrem com a pouca infraestrutura de rede elétrica com que são abastecidas [Tao et al., 2019].

Esse país está situado na zona equatorial do planeta, onde há alta incidência de radiação solar. O índice diário de radiação solar é acima de $4,2KWh/m^2$ na maior parte de seu território [Solargis, 2018]. Na Alemanha, país com a maior capacidade de geração de energia solar fotovoltaica do planeta, há uma média de radiação solar de $1,2kWh/m^2$, muito inferior à encontrada em Burkina Fasso. No entanto, a Alemanha alcançou no verão europeu de 2014 a produção histórica de energia solar fotovoltaica para suprir 50,6% da sua demanda de energia elétrica, mesmo esta sendo alta devido ao alto grau de industrialização do país [Leite, 2018].

Com base nisso, fica evidente o grande potencial de produção de energia fotovoltaica em Burkina Fasso, que acarretaria em um alto impacto sócio-econômico no país. O trabalho de [Tao et al., 2019] explorou a aplicação de técnicas de *Extreme Learning Machine* (ELM) e *Self-Adaptive Evolutionary Extreme Learning Machine* (SaD -ELM) para predição de radiação solar em séries temporais diárias de quatro estações de energia fotovoltaica de Burkina Fasso.

Nesse trabalho foram realizados experimentos com diferentes combinações das variáveis ambientais de temperatura (máxima e mínima), umidade do ar (máxima e mínima), velocidade do vento, deficit de pressão de evaporação e vapor. Neles, foram alcançados resultados melhores que trabalhos que utilizavam técnicas estatísticas como *Auto-Regressive-Integrated-Moving-Average* models (ARIMA) e outros modelos *Artificial Neural Networks* (ANN).

No entanto, os modelos de ELM e SaD-ELM não capturam as relações escondidas entre as variáveis ambientais e padrões temporais ao longo dos dias da série tão eficazmente quanto os modelos baseados em memória de longo ou curto prazo, como o LSTM [Tao et al., 2019]. Os autores sugerem como trabalhos futuros que o uso de técnicas de *Deep Learning* melhorariam ainda mais seus resultados.

O presente trabalho tem como objetivo dar continuidade ao trabalho de [Tao et al., 2019], trazendo melhorias nos resultados de predição da radiação solar através do

uso do modelo LSTM. E, com isso, gerar impacto sócio-econômico positivo na região de Burkina Fasso, ao trazer informações mais acuradas para a instalação e manutenção de painéis solares.

1.1 Trabalhos relacionados

O trabalho de [Zhou et al., 2019] propôs o uso de previsão de séries temporais de produção de energia solar fotovoltaica através de um modelo de Deep Learning. Ele usa duas redes LSTM em séries de temperatura e energia fotovoltaica, respectivamente. Posteriormente, as saídas das duas redes LSTMs entram em uma camada completamente conectada a fim de aumentar a acurácia da predição. Além disso, foi aplicado o mecanismo de atenção, em que os dois LSTMs podem focar nas melhores características de entrada de forma adaptável. Apesar dos bons resultados alcançados, a base de dados utilizada se estende apenas de 2014 à 2018 (da cidade de Shaoxing, província de Zhejiang, China) e contém como variáveis de entrada apenas temperatura e geração de energia.

Já o trabalho de [Arshi et al., 2019] analisou o uso de outras variáveis ambientais, como radiação solar, velocidade do vento e umidade na previsão da geração de energia solar fotovoltaica através de um modelo LSTM. A escolha dos atributos foi feita através da Avaliação de Correlação de Atributos através de *Ranker Search* usando o software Weka ³. A base de dados utilizada foi da localidade de Londres, Inglaterra, de 2013 à 2014.

No trabalho de [Abdel-Nasser and Mahmoud, 2019] é utilizado o modelo LSTM-RNN para previsão de geração de energia fotovoltaica em séries temporais horárias de produção de energia solar com 9000 instâncias de produção de energia solar fotovoltaica da cidade de Aswan e da cidade do Cairo, Egito. Os autores indicam que o uso de variáveis ambientais poderiam melhorar seus resultados.

³O Weka é uma coleção de algoritmos de Machine Learning e Data Mining escrita em Java na Universidade de Waikato, Nova Zelândia

1.2 Organização do trabalho

O trabalho está organizado da seguinte forma, o capítulo 2 oferece referencial teórico para que o leitor tenha maior entendimento do trabalho, o capítulo 3 apresenta a metodologia seguida e o modelo usado para prever a radiação solar das séries temporais de Burkina Fasso, o capítulo 4 apresenta o estudo de caso e análise dos resultados. O capítulo 5 apresenta a conclusão e os trabalhos futuros sugeridos.

2 Fundamentação Teórica

Esse capítulo fornece referencial teórico para que o leitor possa consultar durante a leitura. Como a base de dados consiste de séries temporais, primeiro introduzimos esse assunto. Posteriormente, estudamos os modelos de maneira gradativa, começando com Redes Neurais, *Recurrent Neural Networks*, *Long Short-Term Memory* e Redes Neurais de Aprendizagem Profunda. E ao final, apresentamos breves explicações sobre as métricas utilizadas para avaliação dos resultados neste trabalho.

2.1 Série Temporal

Segundo [Adhikari and Agrawal, 2013], uma série temporal é um conjunto de observações registradas sequencialmente no decorrer do tempo, podendo ser contínuas ou discretas. Em séries contínuas, as observações são medidas em cada instância tempo de maneira não discretizada, por exemplo, leituras de temperaturas ou vazão de rios. Já uma série temporal discreta contém observações medidas em pontos discretos tempo, como medidas da população de uma cidade ou produção de uma fábrica. Em relação à quantidade de variáveis observadas nas medições ao longo do tempo, pode ser considerada univariada, com uma única variável medida, ou multivariada, quando mais de uma variável é observada ao longo do tempo [Adhikari and Agrawal, 2013].

Além disso, as séries temporais possuem componentes determinísticos e não-determinísticos, esses fazem com que não se possa prever com certeza que algo irá ocorrer no futuro. Geralmente, assume-se que elas seguem um modelo estocástico que descreve a distribuição da variável medida [Adhikari and Agrawal, 2013].

Segundo [Adhikari and Agrawal, 2013], as séries temporais, em geral, têm quatro principais componentes:

- **Tendência:** a tendência é o movimento da série a longo prazo. Nesse aspecto elas podem ser crescentes, decrescentes ou estagnadas. O crescimento da população é um exemplo de tendência de crescimento.

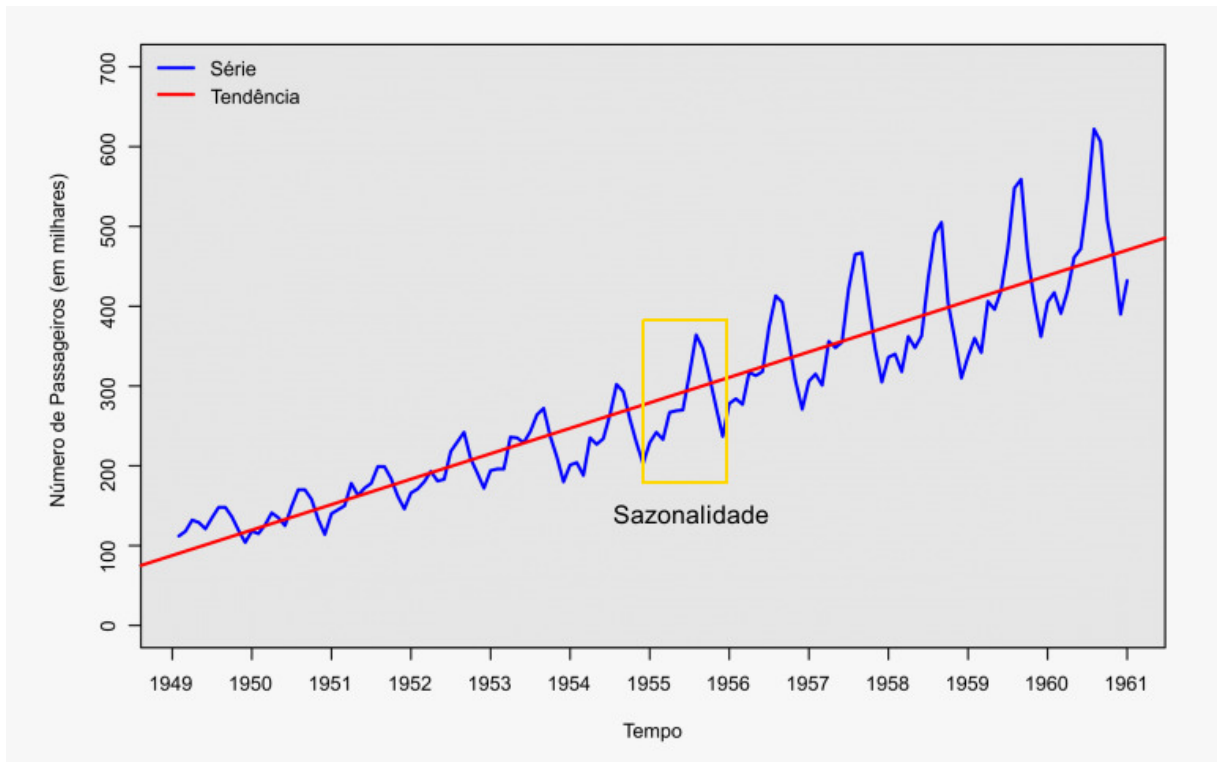


Figura 2.1: Elementos de sazonalidade e tendência da série temporal.

- **Ciclo:** a variação cíclica na série temporal descreve as mudanças de médio prazo, causada por circunstâncias, que se repetem com certa periodicidade. Por exemplo, em séries econômicas ocorrem ciclos começando por prosperidade, declínio, depressão, recuperação, então começando novamente em prosperidade.
- **Sazonalidade:** a sazonalidade pode ser definida como padrões de comportamento que se repetem em específicos intervalos de tempo. A venda de sorvetes aumenta em no verão, a venda de roupas quentes aumenta no inverno, dentre outros exemplos.
- **Irregularidade ou fatores aleatórios:** são causados por influências não previsíveis, que não são regulares e também não se repetem com padrões regulares. Em séries econômicas, por exemplo, a ocorrência de guerras, greves, terremotos influenciam o comportamento da série, mas não podem ser previstos.

Outro conceito importante no estudo de séries temporais é a estacionariedade, uma vez que essa característica é condição necessária para fazer previsões úteis. Ela acontece quando propriedades estatísticas como média e variância independem do tempo. Basicamente, séries temporais que possuem tendência e/ou sazonalidade não são esta-

cionárias e é necessário o uso de técnicas adequadas a previsões que utilizam base de dados que consiste de séries temporais. Na Figura 2.1 evidenciamos um exemplo de uma série temporal com os elementos descritos destacados [Adhikari and Agrawal, 2013].

2.2 Redes Neurais

Segundo [Haykin et al., 2009], redes neurais artificiais foram inspiradas nas redes neurais biológicas. Um neurônio biológico é descrito na Figura 2.2. Através dos dendritos, os neurônios biológicos recebem as entradas, os soma no corpo celular e então produzem a saída pelos axônios, se esta é maior que um limiar – indicada nas sinapses [Meenal and Selvakumar, 2018].

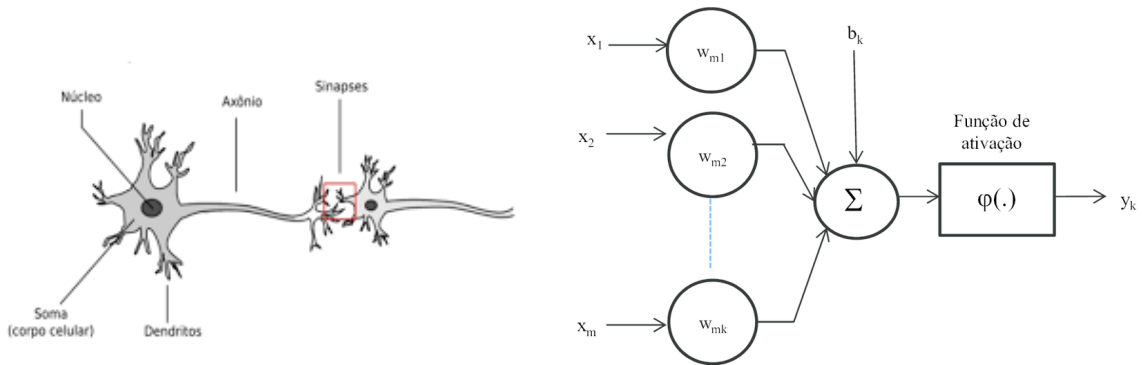


Figura 2.2: Modelo de neurônio biológico e modelo de neurônio proposto em [Haykin et al., 2009].

O elemento básico de uma rede neural artificial (ANN) é o neurônio, descrito na Figura 2.2, que é uma estrutura lógico-matemático que simula o comportamento e funções do neurônio biológico. Os elementos x_1, x_2, \dots, x_n representam as entradas do neurônio artificial e w_1, w_2, \dots, w_n são os pesos - as entradas correspondem aos dendritos e os pesos às sinapses - associados a esses canais de entrada. A informação de entrada é ponderada e processada por uma função de soma e esse novo sinal passa pela função de ativação, que corresponde à diferença de potenciais no núcleo da célula. Assim, os neurônios artificiais são baseados em uma função matemática, que pode ser traduzida na Equação 2.1 [Meenal and Selvakumar, 2018, Haykin et al., 2009].

$$y_i = \varphi_i\left(\sum_{j=1}^m x_j \omega_{i,j} + \theta_i\right) \quad (2.1)$$

onde, x_j é a j -ésima entrada do neurônio i , $w_{i,j}$ é o peso associado, φ_i é a função de ativação, e θ_i é o *bias*, componente que gera variância em relação a entrada.

A função de ativação é geralmente utilizada para adicionar a capacidade de representar mapeamentos funcionais, arbitrários, complexos e não-lineares, entre entradas e saídas. Alguns exemplos de função de ativação são a logística, a tangente hiperbólica (*tangh*) e a *Rectified linear units* (ReLU) [Graves, 2012, Ng, 2017].

Redes neurais podem ser treinadas com o algoritmo *backpropagation*. Ele permite ajustar os pesos de uma rede neural com base no erro (isto é, perda) obtida na iteração anterior. O ajuste adequado dos pesos garante erros mais baixos, tornando o modelo confiável, aumentando sua generalização [Ng, 2017].

2.2.1 Recurrent Neural Networks

Apesar de as Redes Neurais Artificiais (ANN) sem mecanismos de memória apresentarem resultados ótimos, elas têm limitações. A mais notável dessas limitações é a de que elas assumem a independência entre exemplos (instância da base de dados) de teste e treino. Depois que cada exemplo é processado, o estado inteiro da rede é perdida. Se cada exemplo é gerado independentemente, isso não acarreta problemas. No entanto, se eles são relacionados ao longo do tempo, como em vídeos e áudios, isso não modela essas séries de maneira eficaz. Adicionalmente, ANNs geralmente assumem que as instâncias são vetores de tamanho fixo. Então é desejável estender esses mecanismos de aprendizagem para sequências e entradas/saídas de tamanho variável [Lipton et al., 2015].

As Redes Neurais Recorrentes (RNN) têm capacidade de passar dados seletivamente por etapas de sequência, enquanto processam dados da série, um elemento de cada vez. Então elas podem modelar entradas/saídas em que a sequência de elementos é temporalmente dependente [Lipton et al., 2015], [Graves, 2012].

Como mostrado na Figura 2.3, a RNN lê os dados (x) da esquerda para a direita e os parâmetros usados para cada etapa do tempo (W_{ax}) são compartilhados. As conexões horizontais são regidas pelos parâmetros W_{aa} , que são os mesmos para cada instância da série ao longo do tempo. Já os W_{ay} são os parâmetros que influenciam as saídas [Lipton et al., 2015], o que é descrito pela Equação 2.4 [Lipton et al., 2015].

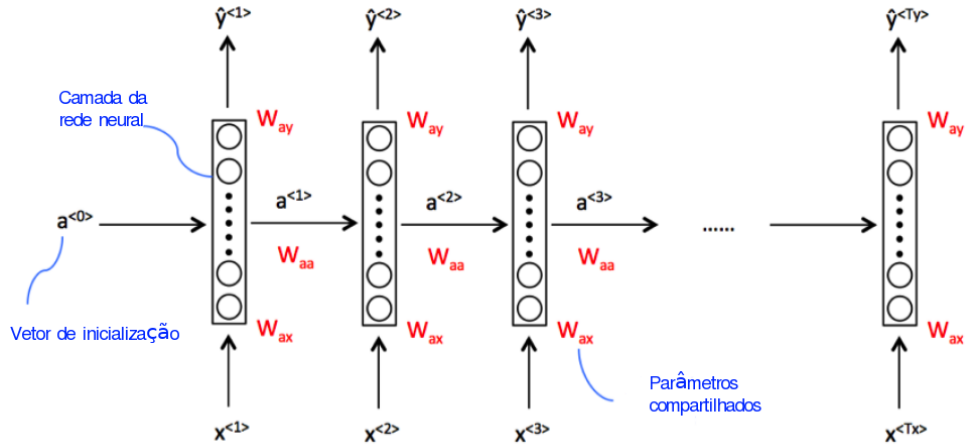


Figura 2.3: Rede neural recorrente. Adaptado de [Cavaioni, 2018].

Uma característica muito significativa que observamos é que, ao fazer uma previsão $\hat{y}^{<i>}$, a rede usa informações não apenas da entrada correspondente $x^{<i>}$, mas também as anteriores - referenciadas pelas Equações 2.2, 2.3. Por exemplo, para prever $y^{<3>}$, obtém informações não apenas correspondentes a $x^{<3>}$, mas também de $x^{<1>}$ e $x^{<2>}$. Além disso, o $a^{<i>}$ representa a saída de uma camada anterior para a próxima camada e $a^{<0>}$ corresponde a inicialização.

$$a^{<0>} = 0 \quad (2.2)$$

$$a^{<t>} = g(W_{aa}a^{<t-1>} + W_{ax}a^{<t>} + b_a) \quad (2.3)$$

$$\hat{y}^{<t>} = g(W_{ay}a^{<t>} + b_y) \quad (2.4)$$

2.2.2 Long Short-Term Memory (LSTM)

O trabalho [Hochreiter and Schmidhuber, 1997] propôs o LSTM com a intenção de superar o problema de *vanishing gradient*. O *vanishing gradient* ocorre, porque ao calcular os erros no processo de *backpropagation* para calibrar os pesos da rede usam-se derivadas para encontrar o quanto houve de mudança em relação a camada anterior. No entanto, quando há um grande número de camadas, ao fazer o processo de *backpropagation*, os erros das últimas camadas interferem pouco nos erros das primeiras camadas. E isso é prejudicial ao aprendizado da rede neural, porque isso significa que os pesos e *bias* das camadas iniciais não serão atualizados de maneira adequada a cada sessão de treinamento.

Como essas camadas iniciais geralmente são cruciais para o reconhecimento de elementos dos dados de entrada, isso pode levar à imprecisão geral de toda a rede [Lipton et al., 2015], [Abdel-Nasser and Mahmoud, 2019],[Ng, 2017].

Para conseguir resolver esse problema, os neurônios das redes tradicionais foram substituídos por células de memória que são conectados através das sucessivas camadas. Cada célula contém portões de entrada, de esquecimento e de saída. O portão de entrada determina como o fluxo de dados de entrada atualizam o estado de memória da célula. O portão de esquecimento controla se a informação deve permanecer na memória ou ser esquecida (e também, o quanto deve ser esquecida). Finalmente o portão de saída diz o quão a saída é afetada pelo valor de entrada e pela memória armazenada [Abdel-Nasser and Mahmoud, 2019].

A Figura 2.4 mostra a célula de um LSTM. Primeiramente, o LSTM tem que decidir qual informação não considerar na célula, o que ele faz através do portão de esquecimento formado pela sigmoide (Equação 2.5). Ele analisa h_{t-1} e x_t , e gera uma saída de números entre 0 e 1 para cada número do estado C_{t-1} da célula, sendo 1 para manter essa informação completamente. Depois, ele decide que nova informação ele vai guardar. Para isso ele faz dois processos: i) o portão de entrada, formado pela camada sigmoide i_t , decide qual valor vai ser atualizado e ii) a função de ativação \tanh cria um vetor de novos candidatos \tilde{C}_t (Equação 2.6), que pode ser adicionado ao estado. Depois disso, multiplica-se o estado antigo de f_t por C_{t-1} , para que sejam esquecidos as informações que se julgou desnecessárias, e multiplica-se i_t por \tilde{C}_t , para manter as informações novas que são úteis (Equação 2.7). Então soma-se o resultado dessas duas multiplicações. Finalmente, decide-se a saída, para isso, a camada sigmoide (Equação 2.8) decide que partes do estado da célula irão para saída, então multiplica-se isso pela \tanh do estado da célula C_t , para que saia apenas as informações que a rede aprendeu que seriam importantes (Equação 2.9).

No LSTM várias dessas células são colocadas em sequência, o que pode ser visto na figura 2.5.

$$f_t = \sigma_t(W_f[h_{t-1}, x_t] + b_i) \quad (2.5)$$

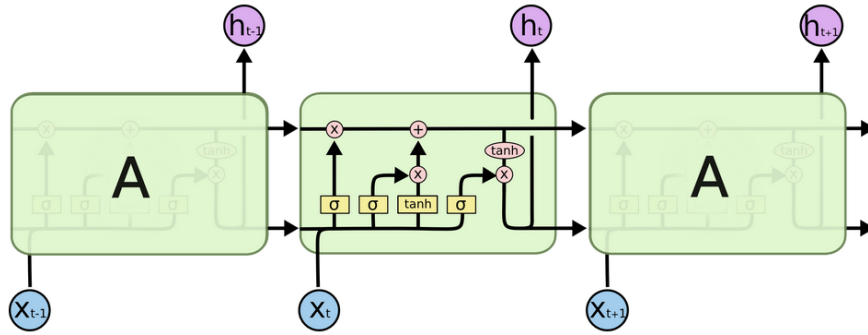


Figura 2.4: As células do LSTM em sequência, adaptado de [Colah, 2018]

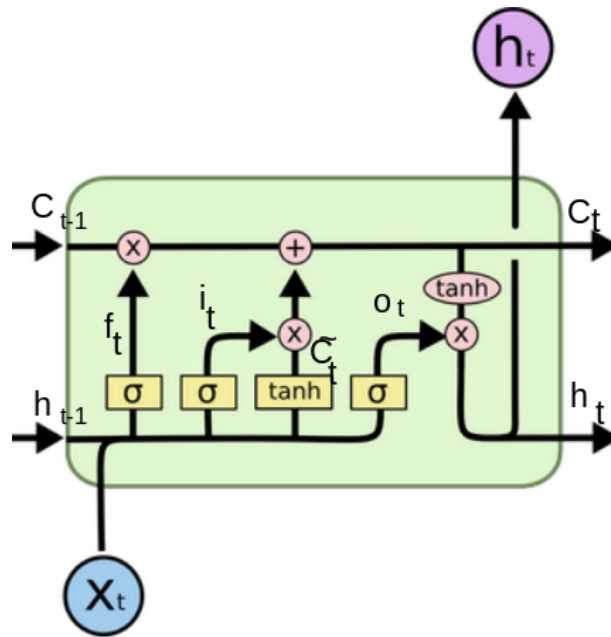


Figura 2.5: Célula do LSTM em detalhe, adaptado de [Colah, 2018]

$$\tilde{C}_t = \tanh((W_C[h_{t-1}, x_t] + b_c)) \quad (2.6)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (2.7)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (2.8)$$

$$h_t = o_t \odot \tanh(C_t) \quad (2.9)$$

2.3 Redes Neurais de Aprendizagem Profunda (Deep Learning)

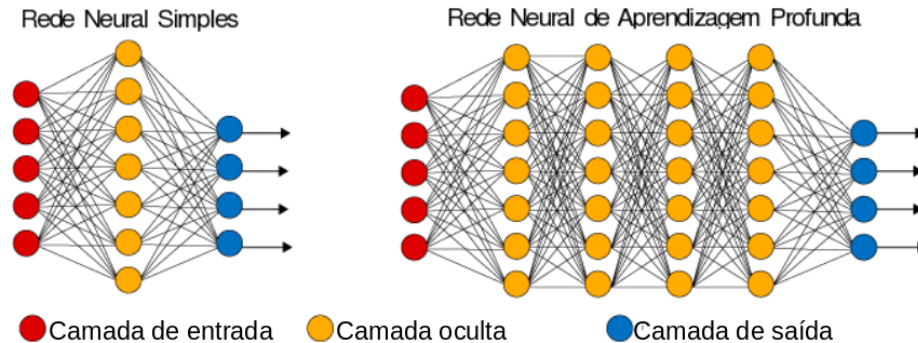


Figura 2.6: Modelos ANN e Modelo de *Deep Learning*, adaptado de [de Engenharia, 2018].

Segundo [Hinton and Salakhutdinov, 2006], *Deep Learning* (DL) pode ser definida como redes neurais artificiais com muitas camadas escondidas, conseguindo capturar relações não-lineares com maior acurácia do que os modelos propostos anteriormente na literatura, como ANNs.

Para atingir esse sucesso, o primeiro passo importante, proposto em [Hinton and Salakhutdinov, 2006, Bengio et al., 2007], foi superar as dificuldades com *vanishing gradient* e fazer treinamentos de DL pela primeira vez. Depois, [Glorot and Bengio, 2010] propôs melhores processos de inicialização e o que tornou possível treinar essas redes com *backpropagation*.

Segundo [Díaz-Vico et al., 2017], outro passo importante foi o uso de processos de regularização, particularmente o *dropout*, em que a rede deixa de considerar alguns neurônios durante seus cálculos, prevenindo assim, que o processo de aprendizagem seja prejudicado, com problemas como *overfitting*. Junto a isso, segundo o mesmo autor, técnicas de processamento de entrada foram introduzidos, como redes convolucionais, proposta por LeCun [Becker et al., 1988], e funções de ativação como a ReLU levaram a mais estabilidade no treinamento.

2.4 Métricas

O trabalho de [Tao et al., 2019] utilizou seis métricas para avaliar o resultado encontrado pelo modelo de previsão de radiação solar: PEARSON-R, VAF RMSE, SI, MAE e NSE.

O PEARSON-R (Equação 2.10) ou Coeficiente de Correlação de Pearson descreve o grau de colinearidade entre o dado medido e o predito. Sua variação é de -1 a 1. Se o índice encontrado for 0, então diz-se que não há relação linear entre a série verdadeira e predita. Apesar de muito usada, essa medida estatística é muito sensível a valores muito extremos e não-sensível a diferenças aditivas ou proporcionais entre a série predita e verdadeira [Moriasi et al., 2007].

$$\text{PEARSON-R} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2.10)$$

Enquanto o índice PEARSON-R é muito sensível a valores extremos, o índice NSE (Equação 2.11) (*Nash Sutcliffe Efficiency* – Eficiência de Nash Sutcliffe) é menos sensível a esses valores visto que considera diferenças quadráticas. Ele determina a magnitude relativa da variância residual (“ruído”) em comparação com a variação dos dados medidos (“informações”). Ele aceita valores de $-\infty$ a 1, sendo 1 o melhor valor. Valores entre 0 e 1 são considerados aceitáveis, enquanto menores que 0 indicam que a média é um melhor valor predito do que o encontrado pelo modelo [Moriasi et al., 2007].

$$\text{NSE} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.11)$$

A VAF (Equação 2.12 (*Variance Accounted For* – Variação contabilizada) indica a variação contabilizada entre duas séries. Se as duas séries são iguais o VAF é 1. Ela é frequentemente usada para verificar a acurácia de um modelo, comparando o dado medido com aquele estimado pelo modelo [TUDEFLLT, 2018, Tao et al., 2019].

$$\text{VAF} = \left(1 - \frac{\text{var}(y - \hat{y})}{\text{var}(y)}\right) \times 100 \quad (2.12)$$

Outra métrica utilizada por [Tao et al., 2019] é o SI (Equação 2.13 (*Scatter Index* – Índice de Dispersão) – que é calculado dividindo-se Raiz do Desvio Quadrático Médio

com a média das observações em cada ponto do plano formado pelos eixos de observações versus dado previsto. Apresenta a diferença de *Root Mean Square* em relação à observação média ou fornece a porcentagem de erro esperado para o parâmetro. Quanto menor o SI, melhor é o desempenho do modelo [Mentaschi et al., 2013].

$$\mathbf{SI} = \frac{RMSE}{\frac{1}{n} \sum_{i=1}^n} \quad (2.13)$$

Como métricas de índice de erro foram utilizados a MAE (*Mean Absolute Error*) e RMSE (*Root Mean Squared Error*). O RMSE (Equação 2.14), muito utilizado na literatura, mede o quanto os dados previstos estão longe da curva formada pelos dados medidos. É interessante, portanto, que o valor de RMSE seja próximo de 0 [Abdel-Nasser and Mahmoud, 2019, Moriasi et al., 2007].

$$\mathbf{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{d_i - f_i}{\sigma_i} \right)^2} \quad (2.14)$$

O MAE (Equação 2.15) mede a média de erros absolutos. Novamente, quanto mais próximo a 0, melhor é o resultado do modelo [Abdel-Nasser and Mahmoud, 2019, Moriasi et al., 2007].

$$\mathbf{MAE} = \frac{1}{n} \sum_{t=1}^n |e_t| \quad (2.15)$$

3 Metodologia

Esse capítulo apresenta a base de dados utilizada, o pré-processamento dos dados e o modelo LSTM utilizado nesse trabalho para previsão de radiação solar em estações fotovoltaicas da região de Burkina Fasso.

3.1 Base de dados

O presente estudo usa a base de dados de quatro estações fotovoltaicas localizadas em Burkina Fasso, país localizado na África Subsaariana pode ser visto na Figura 3.1. Cerca de 70% da produção total de energia elétrica de Burkina Fasso vem de fontes de combustíveis termo-fósseis, o restante vem de hidrelétricas. Em regiões do interior do país, a energia provem de lenha, carvão, resíduos agrícolas e esterco de animais.

Sendo essas fontes baseadas em combustíveis termo-fósseis poluentes e havendo o abastecimento de energia elétrica precário no interior do país, é interessante para o governo explorar fontes de energia renováveis e com capacidade de atender regiões distantes dos grandes centros urbanos [Tao et al., 2019].

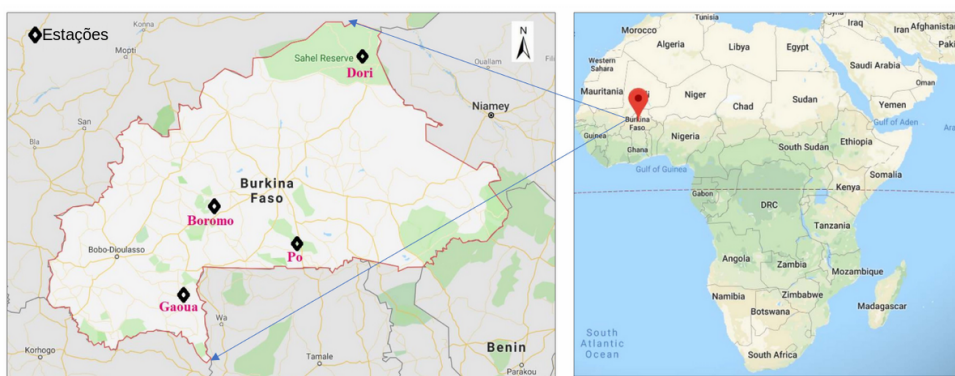


Figura 3.1: Estações de Boromo, Dori, Gaoua e Po, adaptado de [Tao et al., 2019]

A previsão da radiação solar diária foi aplicada como uma variável dependente das seguintes estações: Boromo, Dori, Gaoua e Po. As características ambientais presentes na base são velocidade do vento (WS), temperatura máxima e mínima do tempo (Tmax e Tmin), umidade máxima e mínima (Hmax e Hmin), déficit de pressão de vapor (VPD)

Tabela 3.1: Variáveis ambientais presentes na base.

Variável ambiental	Siglas em inglês	Significados e Unidades SI
WS	Wind Speed	Velocidade do vento, medida em m/s
Tmax	Maximum Temperature	Temperatura máxima, medida em K (Kelvin)
Tmin	Minimum Temperature	Temperatura mínima, medida em K (Kelvin)
Hmax	Maximum Humidity	Umidade máxima, medida em gm/ml^3
Hmin	Minimum Humidity	Umidade mínima, medida em gm/ml^3
VPD	Vapor Pressure Deficits	Déficit de pressão de vapor, medida em kPa
Eo	Evaporation	Evaporação, medida em mm

e evaporação (Eo), resumidas na Tabela 3.1 ⁴. VPD é definido como a umidade máxima do ar quando saturada.

O trabalho de [Tao et al., 2019] investigou oito diferentes combinações das séries de variáveis ambientais como entradas para predizer a radiação solar. O melhor modelo encontrado foi o que envolvia WS, Tmax, Tmin, Hmin, VPD, e Eo. Por isso, utilizaremos esse conjunto no presente trabalho.

As séries temporais utilizadas são diárias de 1/1/1998 a 31/12/2012 (15 anos) ininterruptos, que de 1998 a 2008 (11 anos) foi aplicado a um conjunto de dados de treinamento e de 2008 a 2012 (4 anos) para um conjunto de dados de teste, mesma divisão apresentada por [Tao et al., 2019]. Esses conjuntos de dados de treinamento e teste foram usados para construção e validação de modelos, respectivamente.

3.1.1 Pré-processamento dos dados

A normalização é uma técnica geralmente aplicada como parte da preparação de dados para aprendizado de máquina. O objetivo da normalização é alterar os valores das colunas numéricas no conjunto de dados para usar uma escala única, sem distorcer as diferenças nos intervalos de valores ou perder informações. Os valores dos coeficientes de escala (mínimo e máximo) devem ser calculados no conjunto de dados de treinamento e aplicados para reescalar o conjunto de dados de teste e as previsões. Isso evita contaminar o experimento com conhecimento do conjunto de dados de teste.

As séries das variáveis ambientais, WS, Tmax, Tmin, Hmin, VPD, e Eo, sofreram normalização Min-Max. Nessa abordagem, os dados foram reescalados para um intervalo fixo de 0 a 1, de acordo com a Equação 3.1.

⁴As medidas expressas na tabela correspondem às do Sistema Internacional de Medidas para as grandezas em questão.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (3.1)$$

3.2 Modelo LSTM aplicado na previsão das séries de radiação solar de Burkina Fasso

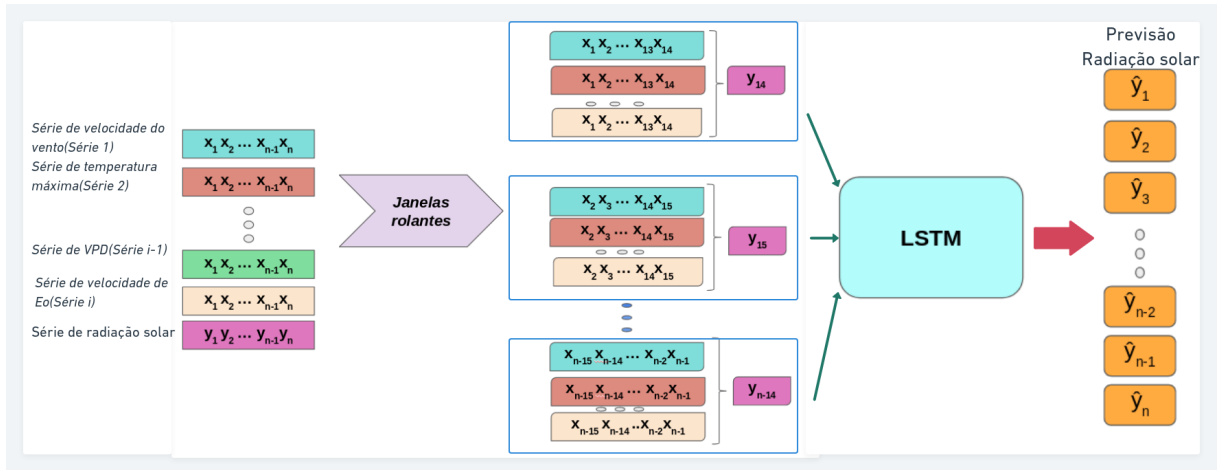


Figura 3.2: Modelo proposto nesse trabalho.

Segundo [Tao et al., 2019] para melhorar a previsão de radiação solar é interessante aplicar um modelo de previsão que consiga capturar o comportamento temporal das séries. Como o modelo LSTM é um tipo especial de RNN capaz de guardar informações de longo prazo, que tem sido amplamente utilizado na literatura em previsão de séries temporais, ele foi o modelo escolhido nesse trabalho para prever as séries de radiação solar a fim de se conseguir um resultado mais confiável do que obtido em [Tao et al., 2019] segundo as métricas avaliadas pelo mesmo.

Desenvolveu-se um modelo com uma camada escondida de LSTM, com uma camada de saída linear, porque ela permite que a saída varie por toda a extensão de valores da variável a ser predita. A inicialização Xavier-He foi utilizada para otimizar os valores de peso. Utilizou-se como otimizador o 'adam', que utiliza pouco recurso computacional e necessita de pouco ajuste de parâmetro. Além disso, são 200 neurônios na camada escondida, como em [Arshi et al., 2019].

As séries de variáveis utilizadas no modelo de previsão foram WS, Tmax, Tmin, Hmin, VPD, e Eo para prever a radiação solar, como mostra a Figura 3.2. No trabalho de

3.2 Modelo LSTM aplicado na previsão das séries de radiação solar de Burkina Fasso 27

[Tao et al., 2019], foi proposto um modelo que utiliza as variáveis ambientais do dia n para prever a radiação solar do dia n . Já no presente trabalho, afim de capturar informações passadas, foi utilizado as variáveis ambientais em uma janela de 14 dias passados para prever a radiação solar do décimo quarto dia. Isso está representado na Figura 3.2 através do uso de um símbolo chaves para expressar que o resultado da radiação solar do dia 14 é uma combinação das variáveis ambientais das séries i dos 14 dias anteriores, por exemplo. Essa janela foi escolhida por experimentos preliminares, começando com janelas 20 dias atrás, como proposto em [Abdel-Nasser and Mahmoud, 2019] para uma série de frequência horária, chegando a uma semana atrás, sendo o resultado de 14 dias o melhor obtido.

4 Estudo de Caso

Neste capítulo aborda-se as configurações do ambiente de experimento, assim como as análises dos experimentos realizados.

4.1 Configurações de Ambiente

Os experimentos foram conduzidos no ambiente Google Colaboratory ⁵, com disponibilidade de 12GB de memória RAM em GPUs. Para a implementação dos modelos foram utilizadas as bibliotecas Tensorflow e Numpy na linguagem Python. E, para análise dos resultados, foram utilizadas as bibliotecas Pandas, Numpy e Matplotlib também na linguagem Python ⁶.

4.2 Análise dos experimentos

As métricas utilizadas para avaliar o resultado do modelo foram: PEARSON-R, VAF, RMSE, SI, MAE e NSE(*Nash Sutcliffe Eficiencia*) como no experimento conduzido em [Tao et al., 2019].

Em relação ao PEARSON-R resultados da Tabela 4.1, para o conjunto de teste, mostram que houve um alto grau de colinearidade entre a série predita e a medida, todas com índices acima de 0.97, exceto para a estação de Dori. Uma explicação para isso é que Dori, localizada no nordeste da região de Burkina Faso, é mais influenciada pelo clima de outros países vizinhos. Portanto, faltam informações mais relacionadas à variabilidade climática. Apesar de o desempenho do modelo ser menor em Dori, ele ainda assim ficou acima do encontrado por [Tao et al., 2019].

Como mostrado na Tabela 4.1, o índice SI foi melhor em todas as estações. Em todas, os resultados do presente trabalho ficou em, no mínimo, 5 magnitudes mais baixo

⁵Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications, acessado em 11/10/2019

⁶Apresentados no trabalho de Python for data analysis: Data wrangling with Pandas, NumPy, and IPython, acessado em 11/10/2019

Métricas	Trabalho	Boromo	Gaoua	Dori	Po
PEARSON-R	Autor	0.983987	0.984423	0.783061	0.979340
	[Tao et al., 2019]	0.982320	0.973168	0.770053	0.954428
VAF	Autor	96.813786	96.883222	60.792374	95.860167
	[Tao et al., 2019]	96.492050	94.652950	59.298030	91.065140
RMSE	Autor	0.125073	0.118768	0.234030	0.134656
	[Tao et al., 2019]	0.722620	0.887638	2.598160	1.175110
SI	Autor	0.000002	0.000002	0.000002	0.000002
	[Tao et al., 2019]	0.033703	0.042528	0.1244820	0.054709
MAE	Autor	0.015655	0.014125	0.054797	0.018156
	[Tao et al., 2019]	0.544310	0.656585	1.980505	0.868936
Nash	Autor	0.964117	0.967054	0.175302	0.955192
	[Tao et al., 2019]	0.962551	0.944803	0.285795	0.900454

Tabela 4.1: Resultados encontrados no conjunto de teste do presente trabalho vs resultados apresentados por [Tao et al., 2019].

do que os apresentados em [Tao et al., 2019]. Em especial, em Gaoua e Po, o índice de dispersão teve grandeza de 10^{-6} , demonstrando que o uso das relações inter-temporais de uma janela de 14 dias passados provoca, de fato, uma melhora muito grande nos resultados de previsão da série temporal de radiação solar.

Observe na tabela 4.1, que para as estações Boromo, Gaoua e Po os valores resultantes de RMSE ficaram abaixo de 0.13, muito abaixo do encontrado em [Tao et al., 2019]. Já em Dori, encontramos 0.234030, apesar de maior que nas outras estações, está muito abaixo de 2.598160 encontrado em [Tao et al., 2019]. O RMSE obteve resultados muito melhores que o encontrado em [Tao et al., 2019] corroborando a importância do uso de um modelo com memória na previsão de radiação solar, como o usado no presente trabalho.

Perceba na Tabela 4.1, Dori, nesse trabalho o resultado foi menor que 0.055, duas magnitudes a menos do que encontrado em [Tao et al., 2019], cujo valor é de aproximadamente 1.98. Já em Boromo, Gaoua e Po, novamente, os resultados foram melhores ficando abaixo de 0.019, enquanto que o melhor encontrado por [Tao et al., 2019] foi, em Boromo, de 0.544310. Novamente, um índice de erro bem estabelecido na literatura mostra que o uso de uma janela de dias passados, em particular a de 14 dias usada nesse trabalho, traz melhoria nos resultados de previsão de radiação solar.

Os resultados relativos ao NSE, nas estações Boromo, Gaoua e Po, são maiores do que os valores encontrados por [Tao et al., 2019]. Os índices encontrados foram maiores

que 0.95, indicando que o valor simulado e o medido estão próximos de uma curva 1:1. No entanto, em Dori esse valor fica em aproximadamente 0.17, um resultado pior do que todos os outros, mas ainda assim no intervalo de 0 a 1, considerado aceitável pela literatura, como pode ser visto na tabela 4.1.

Note na tabela 4.1, que o VAF encontrado no presente trabalho é de aproximadamente 96,81 em Boromo, 96,88 em Gaoua e 95,86 em Po, melhores que em [Tao et al., 2019]. Para a estação de Dori o índice novamente teve valor pior que as demais estações, 60,79, mas ainda assim acima do valor encontrado em [Tao et al., 2019].

Em suma, os resultados encontrados no conjunto de teste, o mesmo utilizado por [Tao et al., 2019], têm SI, MAE e RMSE menores do que os encontrados por ele e PEARSON-R, VAF e NSE maiores do que os encontrados pelo mesmo, exceto para o NSE de Dori. Para Gaoua e Po o modelo foi mais acurado, enquanto que para Dori, o modelo teve resultados menos interessantes se comparado às outras estações, mas ainda assim melhores que [Tao et al., 2019]. É possível ver na Figura 4.2 que em Dori o modelo teve menor capacidade preditiva do que nas estações Boromo, Po e Gaoua, cujas séries de predição acompanham mais fielmente as séries de medições, mostradas na Figuras 4.1 4.3, 4.4.

O principal motivo para obter esse resultados é o de que o LSTM tem capacidade de modelar as mudanças temporais na série, enquanto o método de previsão de [Tao et al., 2019] não utiliza as informações temporais. Em outras palavras, o LSTM com janelamento de 14 dias passados pode capturar conceitos abstratos na série de variáveis ambientais e, assim, houve melhora os resultados da previsão.

Como [Tao et al., 2019] apresentou os resultados do conjunto de treino, apresentamos na Tabela 4.2, os resultados das métricas descritas para o conjunto de treino do nosso trabalho. As análises são análogas às feitas para o conjunto teste. Novamente, Gaoua e Po obtiveram melhores resultados e Dori, piores. É importante notar que os resultados obtidos no conjunto de treino foram melhores dos que os encontrados no mesmo conjunto de treino de [Tao et al., 2019].

Métricas	Trabalho	Boromo	Gaoua	Dori	Po
PEARSON-R	Autor	0.981509	9.807435e-01	0.679370	9.504505e-01
	[Tao et al., 2019]	0.970676	0.963994	0.623858	0.929376
VAF	Autor	96.296188	9.618495e+01	45.874110	8.515733e+01
	[Tao et al., 2019]	94.221080	92.922670	38.836790	86.373670
RMSE	Autor	0.124171	1.236768e-01	0.278206	2.766756e-01
	[Tao et al., 2019]	0.907215	1.074897	3.188808	1.413969
SI	Autor	0.000004	4.409858e-07	0.000006	5.492584e-07
	[Tao et al., 2019]	0.044310	0.053699	0.159305	0.069410
MAE	Autor	0.015418	1.529595e-02	0.077399	7.654936e-02
	[Tao et al., 2019]	0.643799	0.773968	2.226632	1.054562
Nash	Autor	0.962612	9.599267e-01	-0.866303	6.480037e-01
	[Tao et al., 2019]	0.938422	0.922311	-0.700020	0.840394

Tabela 4.2: Resultados encontrados no conjunto de treino do presente trabalho vs resultados apresentados por [Tao et al., 2019].

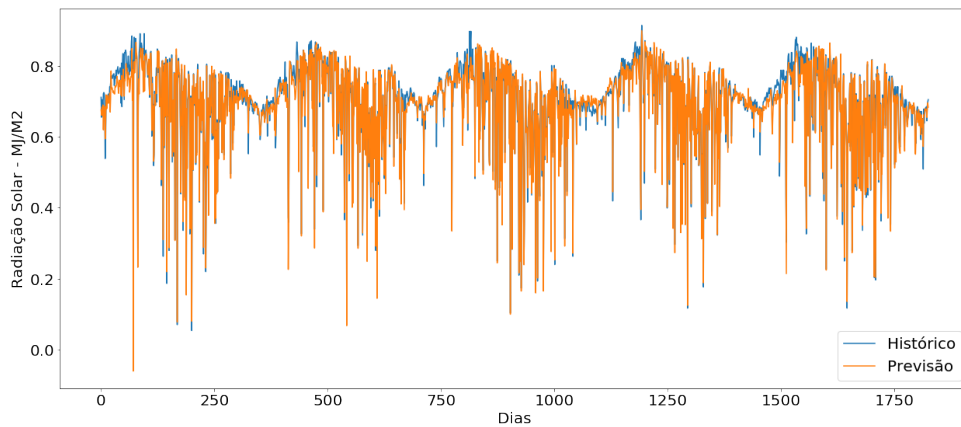


Figura 4.1: Resultado do conjunto teste Boromo.

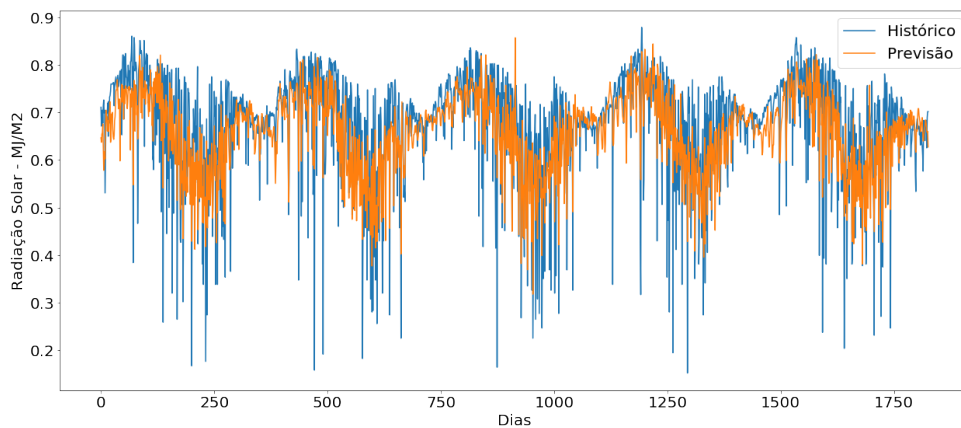


Figura 4.2: Resultado do conjunto teste Dori.

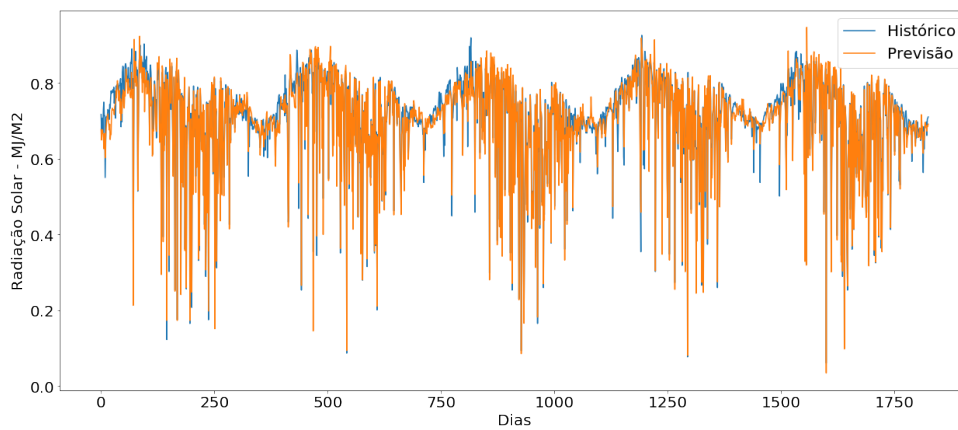


Figura 4.3: Resultado do conjunto teste Po.

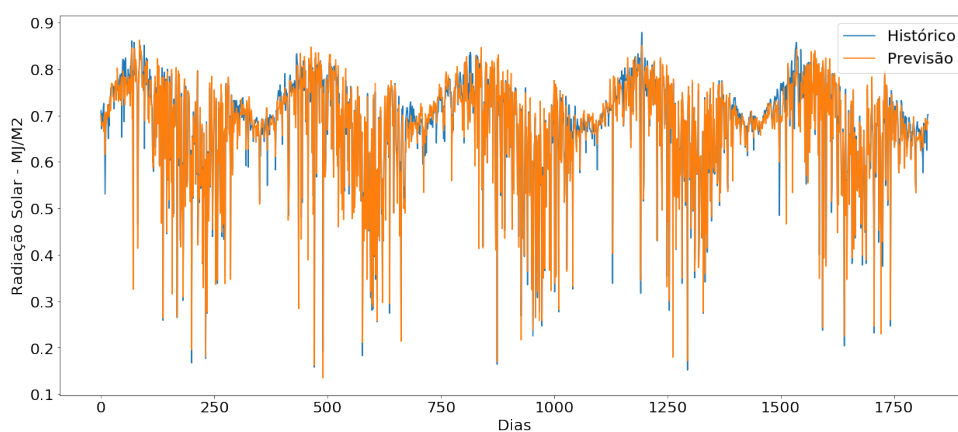


Figura 4.4: Resultado do conjunto teste Gaoua.

5 Conclusão

A previsão de radiação solar através de variáveis ambientais é essencial para instalação de painéis de energia fotovoltaica em lugares em que há pouco recurso financeiro.

[Tao et al., 2019] fez o estudo de modelos de previsão de radiação solar através de variáveis ambientais em uma base de dados de Burkina Fasso que engloba as estações de Gaoua, Po, Boromo e Dori. No entanto, em seu trabalho não foi utilizado um modelo que fizesse uso das relações temporais da série, como sugerido como limitação no próprio [Tao et al., 2019].

Através do uso de um modelo LSTM com janelamento de 14 dias, para as mesmas séries e divisões de conjunto treino/teste, foi possível encontrar valores melhores do que os encontrados em [Tao et al., 2019] utilizando as variáveis ambientais de seu melhor modelo. Isso fica claro através das métricas de VAF, PEARSON-R, SI, NSE, MAE, RMSE, já que em todas elas o modelo LSTM conseguiu melhores resultados.

Esse resultado deixa claro que o LSTM com janelamento de 14 dias passados consegue modelar relações temporais e consegue capturar conceitos abstratos nas séries de variáveis ambientais.

Ainda que o LSTM tenha encontrado resultados melhores que os modelos que não fazem uso de relações temporais, propôs-se como trabalho futuro o uso de modelos de aprendizagem que compartilham os parâmetros da rede, como de transferência de aprendizagem Multitask. Outra melhoria a ser sugerida é o uso de horizontes de predições maiores.

Bibliografia

- Mohamed Abdel-Nasser and Karar Mahmoud. Accurate photovoltaic power forecasting models using deep lstm-rnn. *Neural Computing and Applications*, 31(7):2727–2740, 2019.
- Ratnadip Adhikari and Ramesh K Agrawal. An introductory study on time series modeling and forecasting. *arXiv preprint arXiv:1302.6613*, 2013.
- Sahar Arshi, Li Zhang, and Becky Strachan. Weather based photovoltaic energy generation prediction using lstm networks. 2019.
- Sue Becker, Yann Le Cun, et al. Improving the convergence of back-propagation learning with second order methods. In *Proceedings of the 1988 connectionist models summer school*, pages 29–37, 1988.
- Yoshua Bengio, Pascal Lamblin, Dan Popovici, and Hugo Larochelle. Greedy layer-wise training of deep networks. In *Advances in neural information processing systems*, pages 153–160, 2007.
- Michele Cavaioni. DeepLearning series: Sequence Models, 2018. URL <https://medium.com/machine-learning-bites/deeplearning-series-sequence-models-785>
- Colah. Understanding LSTM Networks, 2018. URL <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>.
- Instituto de Engenharia. “Nobel da Computação” vai para os pais do Deep Learning, 2018. URL <https://www.institutodeengenharia.org.br/>.
- David Díaz-Vico, Alberto Torres-Barrán, Adil Omari, and Jose R Dorronsoro. Deep neural networks for wind and solar energy prediction. *Neural Processing Letters*, 46(3): 829–844, 2017.
- Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feed-forward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256, 2010.
- Alex Graves. Supervised sequence labelling with recurrent neural networks. 2012. URL <http://books.google.com/books>, 2012.
- Simon S Haykin et al. *Neural networks and learning machines/Simon Haykin*. New York: Prentice Hall,, 2009.
- Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. *science*, 313(5786):504–507, 2006.
- Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.
- Solar Energy Industries Association. Photovoltaics, 2019. URL <https://www.seia.org/initiatives/photovoltaics>.

- Waldir Leite. Solar resource maps of Burkina Faso, 2018. URL <https://www.proifes.org.br/artigos/sustentabilidade-energetica/>.
- Zachary C Lipton, John Berkowitz, and Charles Elkan. A critical review of recurrent neural networks for sequence learning. *arXiv preprint arXiv:1506.00019*, 2015.
- R Meenal and A Immanuel Selvakumar. Assessment of svm, empirical and ann based solar radiation prediction models with most influencing input parameters. *Renewable Energy*, 121:324–343, 2018.
- Lorenzo Mentaschi, Giovanni Besio, Federico Cassola, and A Mazzino. Problems in rmse-based wave model validations. *Ocean Modelling*, 72:53–58, 2013.
- Daniel N Moriasi, Jeffrey G Arnold, Michael W Van Liew, Ronald L Bingner, R Daren Harmel, and Tamie L Veith. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, 50(3):885–900, 2007.
- Andrew Ng. Neural networks and deep learning. *Coursera*. <https://www.coursera.org/learn/neuralnetworks-deep-learning>, 325, 2017.
- Solargis. Solar resource maps of Burkina Faso, 2018. URL <https://solargis.com/maps-and-gis-data/download/burkina-faso>.
- Hai Tao, Isa Ebtelahj, Hossein Bonakdari, Salim Heddami, Cyril Voyant, Nadhir Al-Ansari, Ravinesh Deo, and Zaher Mundher Yaseen. Designing a new data intelligence model for global solar radiation prediction: Application of multivariate modeling scheme. *Energies*, 12(7):1365, 2019.
- TUDEFLT. VAF, 2018. URL <https://www.dcsc.tudelft.nl/~jwvanwingerden/lti/doc/html/vaf>
- Kejun Wang, Xiaoxia Qi, and Hongda Liu. Photovoltaic power forecasting based lstm-convolutional network. *Energy*, page 116225, 2019.
- Hangxia Zhou, Yujin Zhang, Lingfan Yang, Qian Liu, Ke Yan, and Yang Du. Short-term photovoltaic power forecasting based on long short term memory neural network and attention mechanism. *IEEE Access*, 7:78063–78074, 2019.