

# Desenvolvimento de um modelo em python de predição em rede neural para dados de LIDAR e CRDS

Pedro Amaral Medeiros e Eduardo Landulfo  
Instituto de Pesquisas Energéticas e Nucleares(IPEN)

## INTRODUÇÃO

Este projeto visa desenvolver um modelo de predição usando métodos de rede neural como RNN(Recurrent Neural Network) utilizando a linguagem python para os dados de LIDAR e CRDS.

Essa linguagem tem o diferencial de ser mais simples e ter diversas bibliotecas para efetuar cálculos de forma eficiente disponibilizadas pela comunidade. Levando em conta a qualidade e quantidade de dados, o CRDS é uma boa escolha para aplicar modelos de predição com aprendizado de máquina tendo os dados do LIDAR como apoio ao modelo.

Há um grande interesse em estudos dos processos de como as partículas de aerossóis interferem no sistema climático em grandes metrópoles, regiões poluídas e povoadas, buscando a compreensão destes, impactando não só à qualidade do ar, mas também a interferência gerada no clima local e/ou regional [1].

## OBJETIVO

O objetivo da iniciação é o desenvolvimento de um modelo de predição para os dados de LIDAR e CRDS usando metodologia de rede neural com a linguagem python..

## METODOLOGIA

Neste trabalho foi utilizada a linguagem python para o modelo de predição de dados do LIDAR e CRDS e para isso foram utilizadas algumas bibliotecas da linguagem python como:numpy, scipy, pandas ,scikit-learn, keras, tensor-flow. Para a construção da rede neural foi usado o

modelo de RNN (Recurrent Neural Network)usando camadas LSTM (Long short-term memory) [2].

Primeiramente foi feito um estudo sobre a correlação dos dados disponíveis do CRDS (CH<sub>4</sub>,CO<sub>2</sub>,CO,H<sub>2</sub>O) para escolher os dados que seriam úteis para treinar o modelo considerando que dados com alta correlação podem diminuir o erro do modelo.

Após esse estudo, o modelo de rede neural foi construído na seguinte ordem:Abertura do arquivo de dados do CRDS;transformar a escala dos dados para o intervalo de 0 a 1; separar os dados para treino, teste e validação[3].

Com isso, foram feitas predições de 24 horas variando a quantidade de camadas LSTM para observar a melhora nas predições usando o modelo com quatro camadas.

## RESULTADOS

Para entender a relação entre os dados do CRDS,nesse caso, CO e CO<sub>2</sub> primeiro foi feito o plot das duas concentrações(ppm) para deixar claro a correlação.

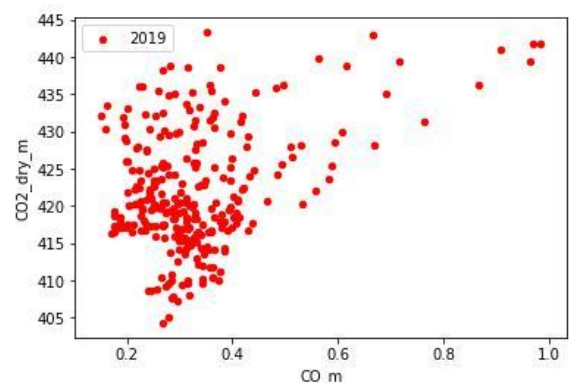


Figura 01: gráfico da concentração de CO2 e CO em ppm de 2019

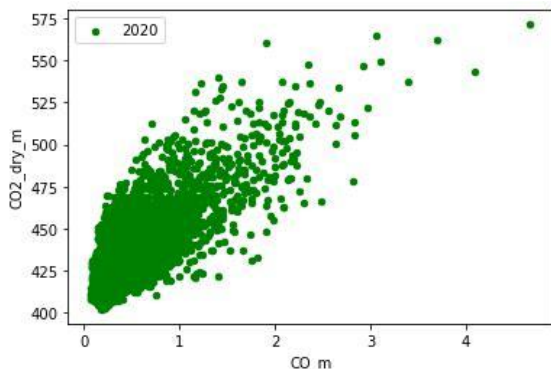


Figura 02: gráfico da concentração de CO2 e CO em ppm de 2020

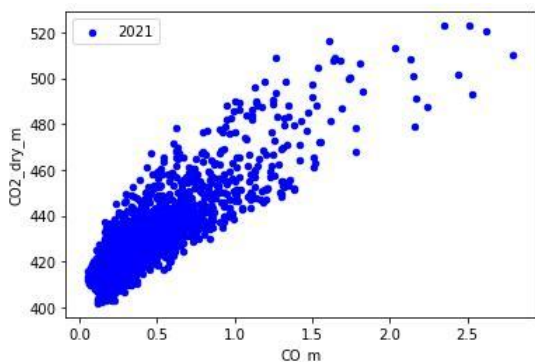


Figura 03: gráfico da concentração de CO2 e CO em ppm de 2021

Após a análise das correlações foi escolhido o ano de 2021 para o treino do modelo

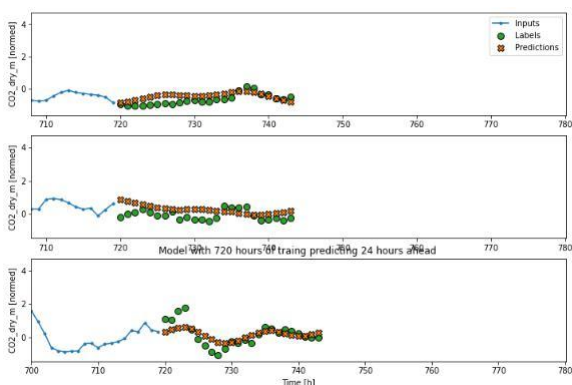


Figura 04: gráfico da das previsões com 720 de treino para o modelo com uma camada LSTM em que a diferença de tempo entre as três previsões é de 100 horas.

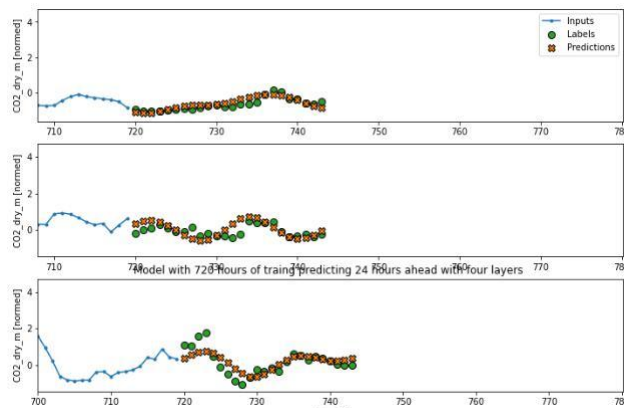


Figura 05: gráfico da das previsões com 720 de treino para o modelo com quatro camadas LSTM em que a diferença de tempo entre as três previsões é de 100 horas.

## CONCLUSÕES

O modelo foi construído atingindo resultados satisfatórios considerando sua primeira aplicação, também foi feita com sucesso a comparação entre modelos treinados com quantidades diferentes de camadas LSTM.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [1] LANDULFO, E. ; PATRÍCIA FERRINI RODRIGUES; FÁBIO JULIANO DA SILVA LOPES; RIAD BOYAROU. Measurement of tropospheric aerosol in São Paulo area using a new upgraded Raman LIDAR system. (E. I. Kassianov et al.,Eds.)1\_nov.\_2012\_Disponível em:<<http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.975060>>
- [2] Pawar K., Jalem R.S., Tiwari V. (2019) Stock Market Price Prediction Using LSTM RNN. In: Rathore V., Worring M., Mishra D., Joshi A., Maheshwari S. (eds) Emerging Trends in Expert Applications and Security. Advances in Intelligent Systems and Computing, vol 841. Springer, Singapore. [https://doi.org/10.1007/978-981-13-2285-3\\_58](https://doi.org/10.1007/978-981-13-2285-3_58).
- [3] [https://www.tensorflow.org/tutorials/structured\\_data/time\\_series#the\\_weather\\_dataset](https://www.tensorflow.org/tutorials/structured_data/time_series#the_weather_dataset)

## APOIO FINANCEIRO AO PROJETO

CNPq