Analyzing COVID-19 Data By Means of Fast Fourier Transforms Through Autoencoders

Christopher Torres, Jesse Ibarra, Gil Gallegos
New Mexico Highlands University, National Science Foundation,
Computer/Mathematical Sciences (CMS)

Overview

- SARS-CoV-2 (COVID-19) pandemic has created extremely large and complex data sets (availability presents its own challenges).
- Distilling dense-information rich transformed data.
- Seeking fruitful insight from the salient features of the data.
- Reduce the complexity (dimensionality) of the data while maintaining the integrity of original data using machine learning autoencoder methods.
- Using an autoencoding process and mathematical modeling by means of Fast Fourier Transform and low pass filtering.

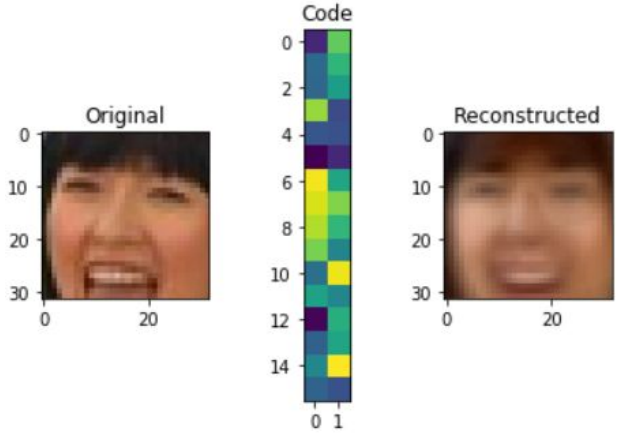
Related Work

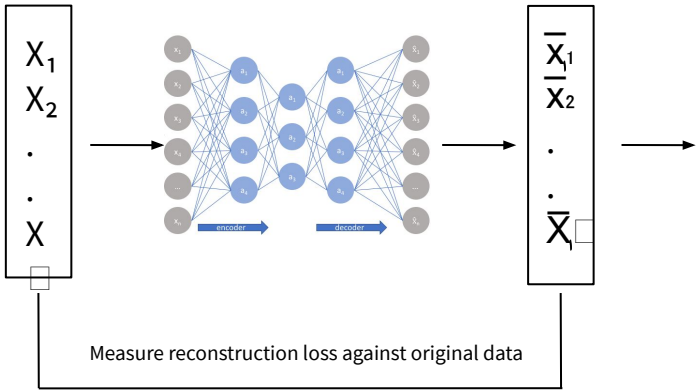
- In a related study in the Journal of Robotics, Networking and Artificial Life, Yokkampon, Umaporn, et al. proposes an autoencoder method with spiking raw data to the frequency domain.
    - Analyze and predict the anomaly case among the standard data set and compare with original data (data set was real-world data from factory automation).
    - The combination of frequency domain and original data has the potential to improve the validity and accuracy in detecting anomalies.
    - Concluded that anomaly detection improves accuracy. Therefore, analyzing time-series data in the frequency domain can be efficient in detecting anomalies for time series data.

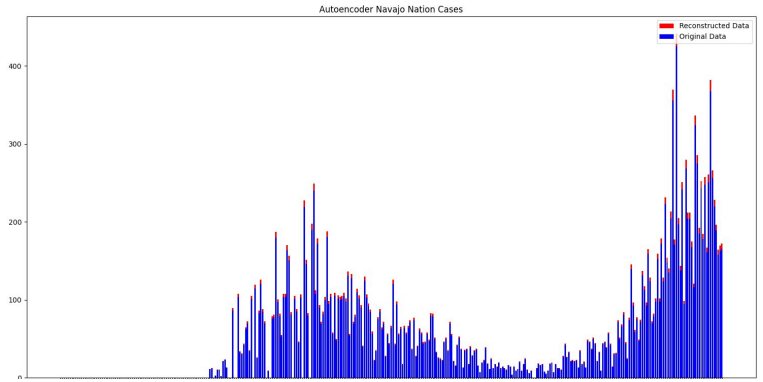
Image source: https://stackabuse.com/autoencoders-for-image-reconstruction-in-python-and-keras/

“What Is An Autoencoder?”

An autoencoder is an **unsupervised artificial neural network** that learns how to compress and encode data. It does this to “learn” how to reconstruct the data back from the **reduced encoded representation** to a representation that is as close to the original input as possible. It works in two key parts: **encoder** and **decoder**.

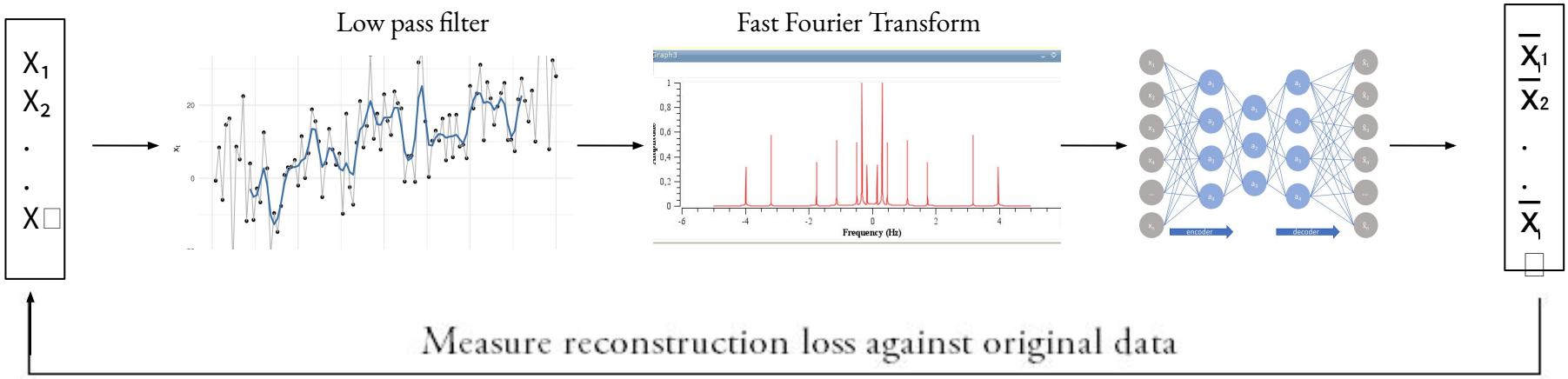




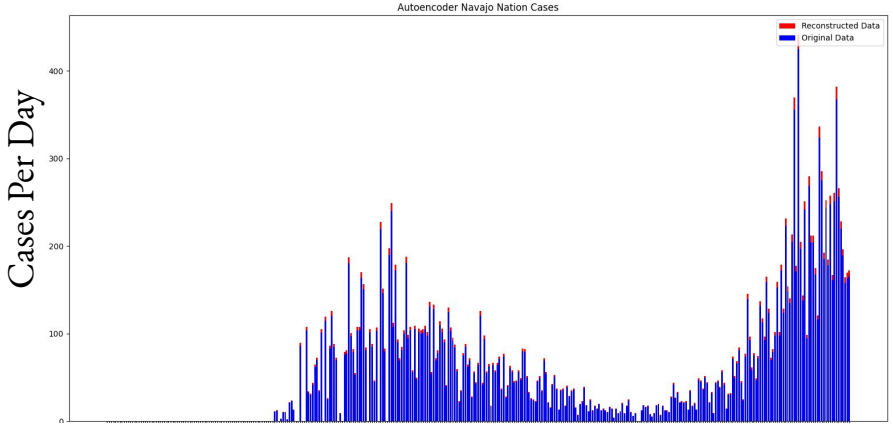


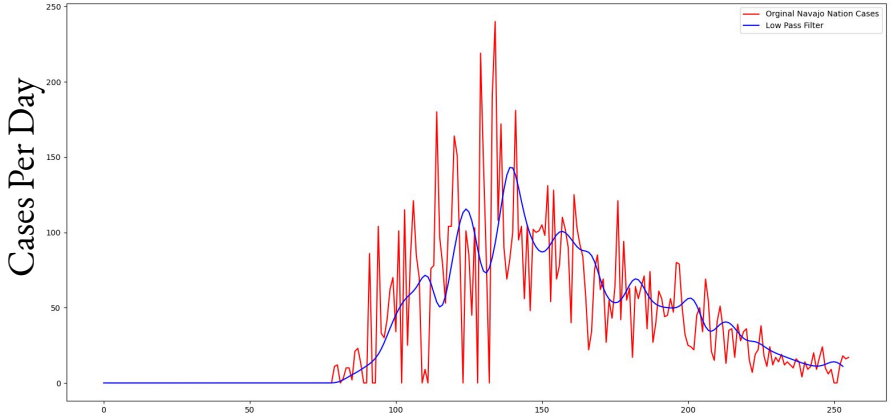
Methods

- Using an autoencoder neural network, the data is deconstructed, compressed, and reconstructed
- This provides the neural network a way to **learn** a representation of the data.
- Using **low pass filters** distills and removes **noise** from the original dataset.
- Utilizing **Fast Fourier Transform** enables the conversion into a **frequency domain** so that the data can be further analyzed.



Observing the Deconstruction and Reconstruction of the Original Data

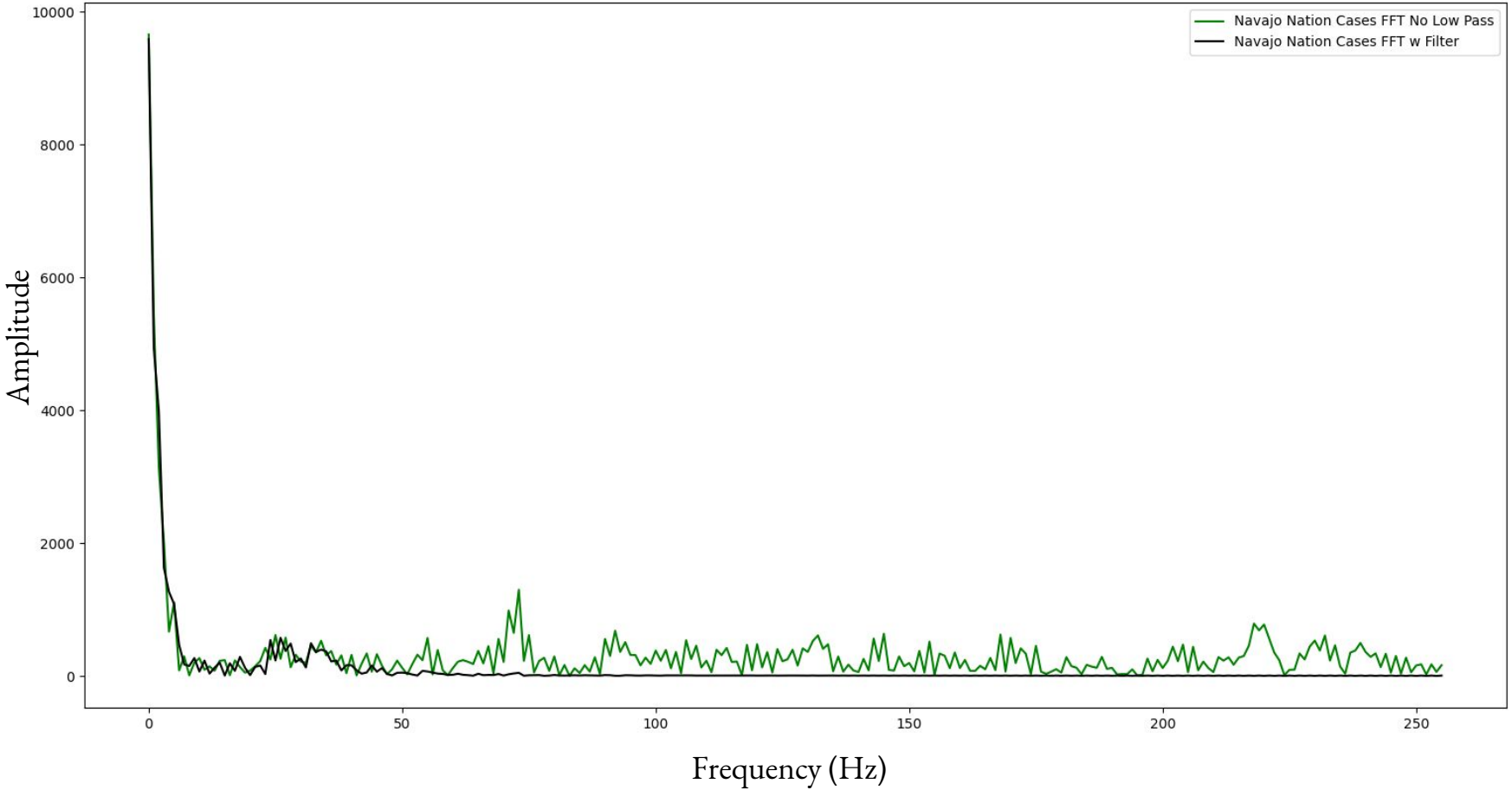




<<<Iterations by day>>>

- The data in this study was collected by day from the Navajo Nation's Dept. of Health
- Reconstructing the original data through the AE
- Passing the original data through a low pass filter to distill the raw data.
- Transform the filtered data to a frequency domain using a Fast Fourier Transform.

Transforming The Data



Conclusions

- Thus far the autoencoder was able to deconstruct and reconstruct the original data set and learn a representation of data.

- The data was able to be passed through a low pass filter and then converted to a frequency domain (Fast Fourier Transform).

- We have been able to reduce the dimensionality of the dataset, extracting real information from the complex domain.

- Observation and analysis of the frequency state is still in process before converting the data back to the time and space domain.

Future Work

- Convert the reconstructed frequency state data back to the time and space domain. (conversion not reversion).

- Combine with original data with the Fast Fourier Transform values.

- Develop and model a metric, as did Yokkampon, Umaporn, et al., for performance evaluation of the accuracy of anomaly detection. Detecting anomalies will be instrumental in testing policy and what was working and not working over the span of the time domain.

- Determine that the data is suitable for forecasting (SIR analysis, QSIR, etc.)

References

- Abdelaal, Ali. “Autoencoders for Image Reconstruction in Python and Keras.” *Stack Abuse*, Stack Abuse, stackabuse.com/autoencoders-for-image-reconstruction-in-python-and-keras/.

- “New COVID-19 Cases & Deaths By Day on Navajo Nation.” *Infogram*, infogram.com/new-covid-19-cases-and-deaths-by-day-on-navajo-nation-1hke60rr5ex365r.

- Peng, Roger D. “A Very Short Course on Time Series Analysis.” *4.2 Filtering Time Series*, 2 Apr. 2020, bookdown.org/rdpeng/timeseriesbook/filtering-time-series.html.

- The SciPy community. “Fourier Transforms (Scipy.fft).” *Fourier Transforms (Scipy.fft) - SciPy v1.6.1 Reference Guide*, docs.scipy.org/doc/scipy/reference/tutorial/fft.html.

- Yokkampon, Umaporn, et al. "Autoencoder with Spiking in Frequency Domain for Anomaly Detection of Uncertainty Event." *Journal of Robotics, Networking and Artificial Life* 6.4 (2020): 231-234.

- Zhang, Chunkai, et al. "VELC: A new variational autoencoder based model for time series anomaly detection." *arXiv preprint arXiv:1907.01702* (2019).

Acknowledgements

- NSF DMR 2031548 (RAPID)
- NSF DMR 1523611 (PREM)

NEW MEXICO HIGHLANDS UNIVERSITY

NSF