Lightscline
Data Reduction AI

Analyzing acoustic data using 10x lighter/faster SwaP-C AI

March 2024
Lightscline, State College, PA, USA

About Lightscline

At Lightscline, we are redefining the 1940s principles of scientific computing (Shannon Nyquist theorem) using AI. Lightscline AI's 4 lines of code reduces 90% of sensor data infra & human costs by training proprietary machine learning models to selectively focus on just 10% of the raw data.

Our selective sensing extends end-to-end in a machine learning deployment scenario - data collection through prediction - using just 4 lines of code. This means that our neural networks give same accuracies while being >10x faster and energy efficient, as they only analyze 10% of the raw data.

Value props for acoustic signal applications:

1. >10x reduction in edge computing power than conventional approaches
2. >10x faster for same analytics accuracies
3. 90% reduction in data infra costs
4. 10x productivity increase (data scientists/machine learning/embedded engineers - more deployment)
5. Helps data scientists select 100 useful windows from 10MM+

Use-cases:

1. **Preserving aircraft take-off signal using just 10% of the raw acoustic data**

In this use-case, we show that just 10% of the raw acoustic data is sufficient to preserve the signal and important phenomena like Doppler shift in an aircraft takeoff signal. The recovered spectrogram shows that Doppler shifts for non-stationary temporal signals can be calculated using a reduced number of measurements. The reconstructed spectrogram can be used for calculating the engine speed, rate of change of frequencies, velocity, and height of the aircraft at take-off. Spectrograms and similar techniques are also useful in marine applications including object detection, marine life detection, etc.

Figure 1 shows the aircraft take-off sound signal and the spectrogram of a 14 second interval of the aircraft take-off sound signal. Figure 2 shows the reconstructed spectrogram using just 10% of the raw data, while showing the Doppler shifted region and frequencies.

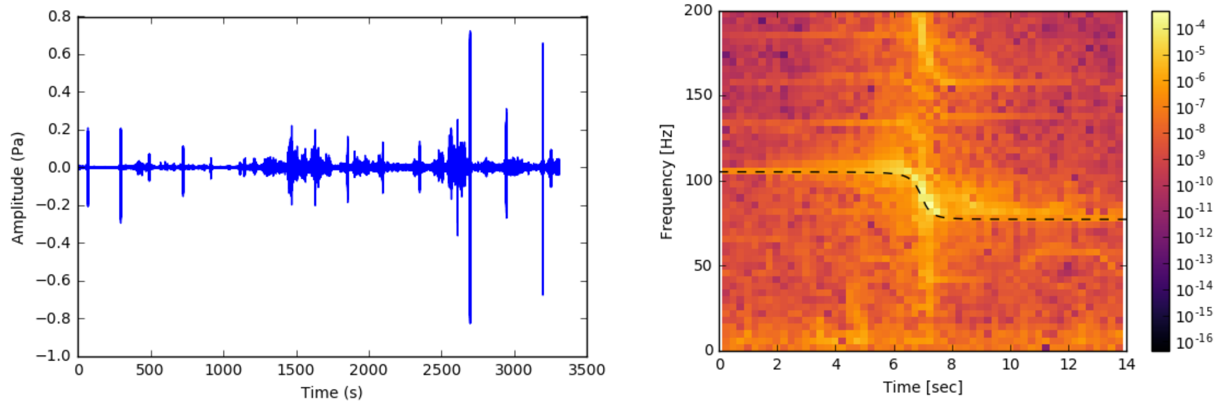

Fig.1. (a) Aircraft take-off sound signal (b) Spectrogram of a 14 second interval of the aircraft take-off sound signal.

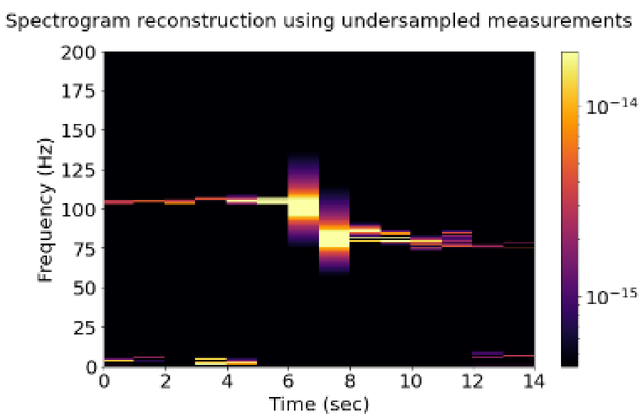

Fig.2. Reconstructed spectrogram using just 10% of the raw data focusing on the Doppler shifted take-off section

You can read more here.

2. Preserving calibration signal using just 10% of the raw acoustic data

Acoustic/sound signals carry important information about the physics of the process and are used in a variety of fault diagnosis and predictive analytics applications. However, due to the high sampling frequencies involved, the amount of data collected for analysis can quickly reach ~10 GB / day / sensor. With 100s of sensors deployed across 1000s of assets, we're suddenly dealing with TBs of sensor data. Lightscline AI exploits the redundancy in real world sensor data to selective focus on the 10% important data, leading to a 90% reduction in the infrastructure and human related time and costs involved in making sense of all this data.

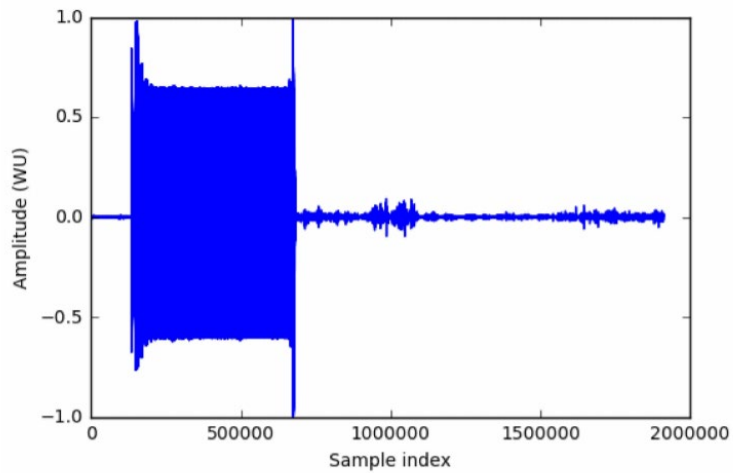

Fig.3. Sound calibration signal

Now, we will look at the FFT made from all the raw data and compare it with the FFT made from just 10% of the raw data.

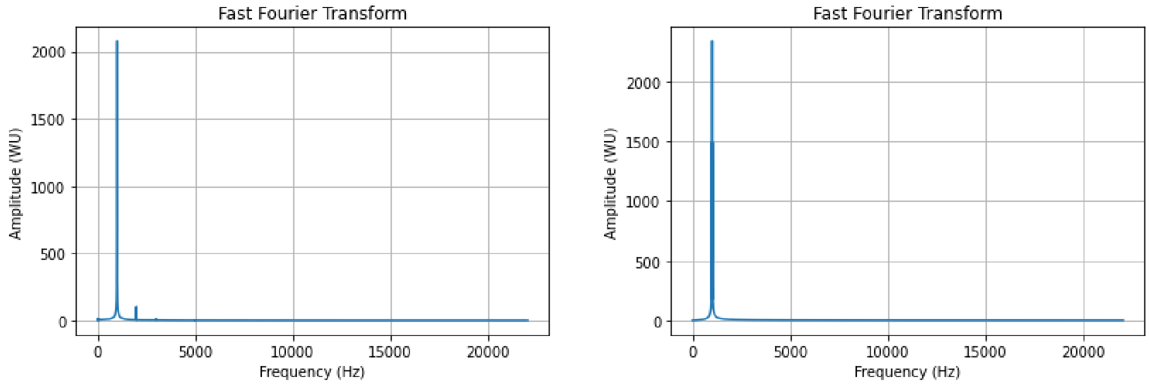

Fig.4. FFT made from all the raw data vs FFT made with just 10% of the raw data

Moreover, FFT is based on the 1940s Nyquist theorem, which requires collection of all the raw data resolved at Nyquist rates. Lightscline's AI goes a step further by not collecting all data points, but learning the amount of data needed for different tasks by using a training process. From a user perspective, this all happens using just 4 lines of code, can be setup within 10 minutes, and does not need any external data sharing. Figure 5 shows how to get started with Lightscline AI using just 4 lines of code.

Lightscline AI

```
from lightscline.lightscline import LightsclineEdge
## Load data into Lightscline
ls = Lightscline(data=data,fs = SAMPLING_FREQUENCY)
## Reduce the amount of data by 70% of the original
ls.reduce_and_preprocess_data(per_reduction=70)
## Train the model
ls.train_model(verbose=True,n_iters = 1000)
## checking the results
ls.test_model()
```

- 4 lines of code to get started
- Setup within 10 mins
- No data sharing required

Fig.5. Lightscline AI – get started using 4 lines of code

Lightscline AI is 10-74x more efficient than a standard neural network while using 100x less data. These tests were conducted on the Case Western Reserve University bearing fault dataset. We observe linear performance improvements like computational complexity in the power, storage, transmission, and latency requirements by using Lightscline AI.

Conclusion

Using just 4 lines of code, Lightscline AI performs end-to-end data collection through prediction with >10x speed and energy efficiency over conventional approaches governed by the Shannon-Nyquist sampling theorem. This leads to orders of magnitude savings in (i) data infrastructure costs and time, and (ii) human resource efficiency. Finally, this enables several new applications not possible today due to extreme SWaP-C requirements of real-time embedded compute. End users can easily validate for themselves using a free trial by going here. Try Lightscline's AI now.

Reach out to info@lightscline.com for any queries.