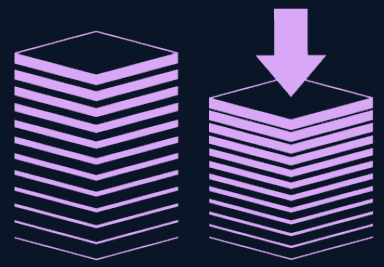


Lightscline

Data Reduction AI

Leveraging AI to generate same benefits from 10% data

Problem

Sensor data

73+ Trillion GB by 2025

Infrastructure cost

$100M+ annual cloud costs

Human capital cost

1 hour of data = **40+ hours** of analysis

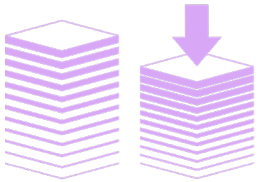
Lightscline
Data Reduction AI

Solution

Leverage AI to automatically identify the 10% high-value data from voluminous data streams

- Reduces 90% of the AI infra and people related time & costs
- Enables new embedded AI applications currently unobtainable due compute, latency, memory, & time bottlenecks

Product



LS Compute

100x less infrastructure costs for training and inference

25k lines of patented code + 5 years R&D

into 4 lines of user-facing code

```
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()
```

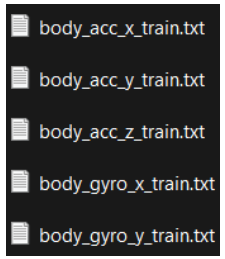
- Setup within 10 mins
- On-prem / cloud hosting
- No data sharing required

Conventional vs. Lightscline AI workflow

Conventional techniques

>40 hours

Raw data files



Manual feature extraction

```
555 angle(tBodyAccMean,gravity)
556 angle(tBodyAccJerkMean),gravityMean)
557 angle(tBodyGyroMean,gravityMean)
558 angle(tBodyGyroJerkMean,gravityMean)
559 angle(X,gravityMean)
560 angle(Y,gravityMean)
561 angle(Z,gravityMean)
```

Model development

- SVM
- KNN
- ANN..

Metrics

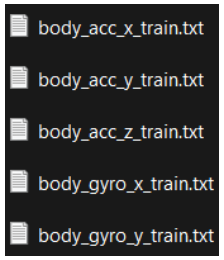
97% accuracy

Lightscline AI

~1 hour

10x less compute & infra costs

Raw data files



Lightscline workflow

```python
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()
```

Metrics

97% accuracy

>10x model size reduction

>10x faster

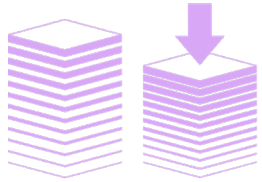
Lightscline
Data Reduction AI

Two more things..

Our end-to-end offering

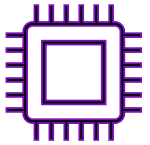
LS Data Labeling (DL)

10000x faster, cheaper than traditional manual labeling



LS Compute

100x less infrastructure costs for training and inference



LS Edge

Deploy models to your hardware

Customer case-studies

LS DL

Enterprise energy monitoring

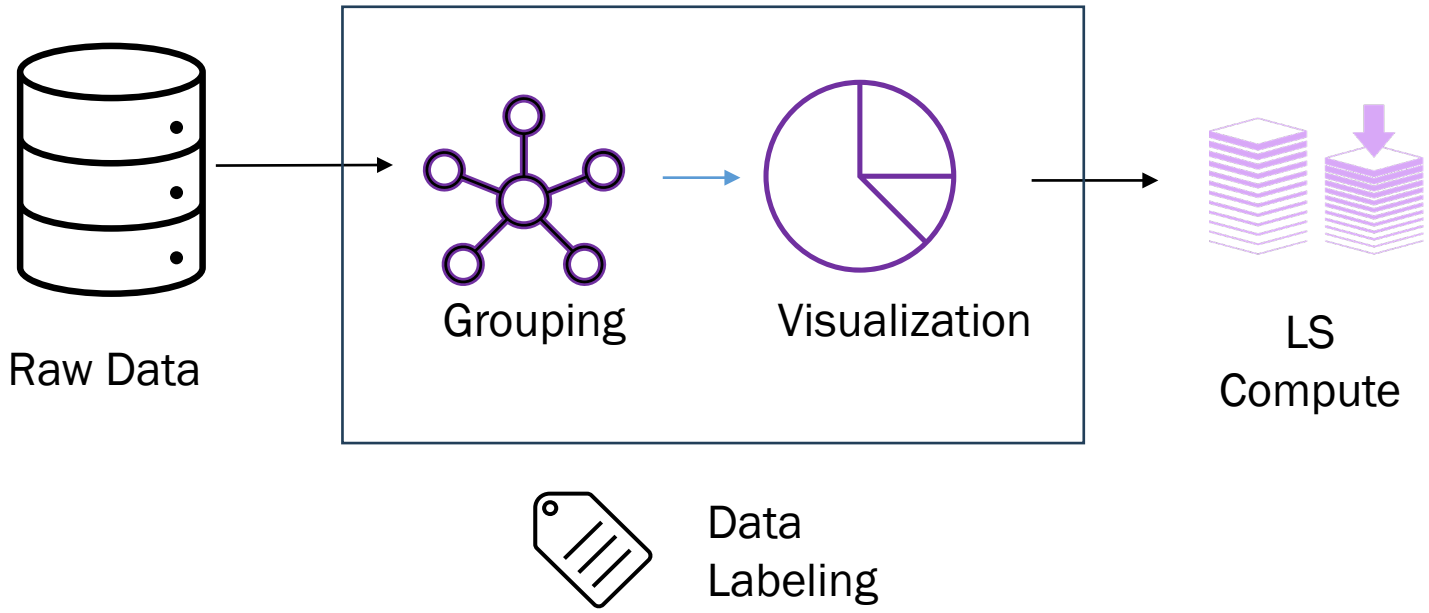
LS Compute

Asset Health Management

LS Edge

Human Activity Recognition

Lightscline Data Labeling



Enterprise energy monitoring

Helps data scientists select 100 useful windows from 10MM+ collected over 3 years

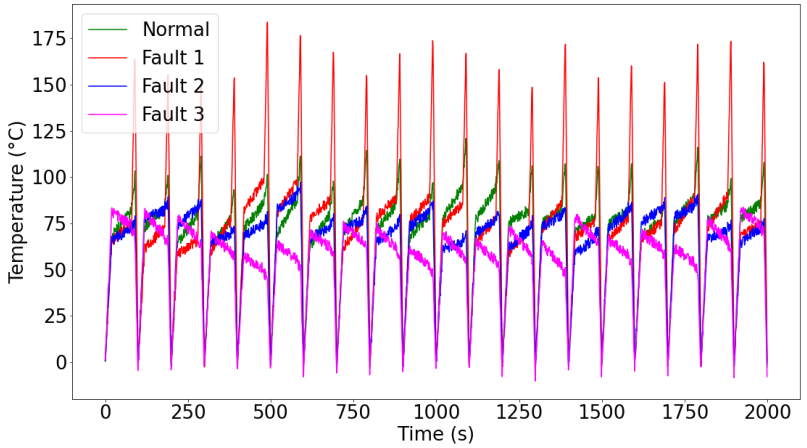
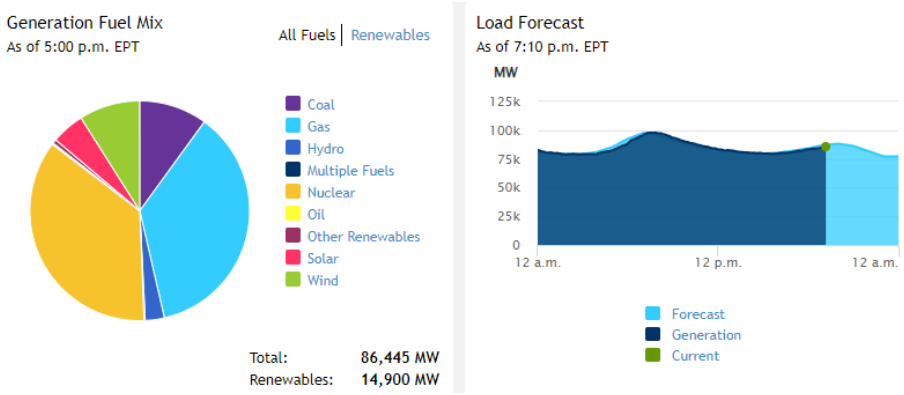


Figure out the reason for 100 MWh excess energy consumption



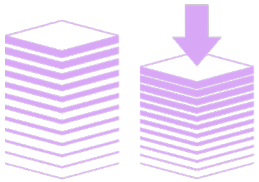
Analyze 100 useful windows which were previously hidden in an S3 bucket

% data used	Accuracy
8	99.9
10	99.9
15	100

*enables applications not possible today

Lightscline
Data Reduction AI

Lightscline Compute



LS Compute

100x less infrastructure costs for training and inference

Asset Health Management

BOSCH

We tried the free version on other datasets and accuracies are mentioned in the below table:

S No	Dataset	Sampling Frequency	Accuracy
1.	Bearing	97656 Hz	86%
2.	Wind turbine planetary gearbox	40000 Hz	100%
3.	Unbalanced Impeller Centrifugal Blower	8000	95.2%
4.	Loose and Dent Impeller Centrifugal Blower	8000	84.6%

With just 10% data

Asset Health Management

BOSCH

From
SX/BSV-TH1

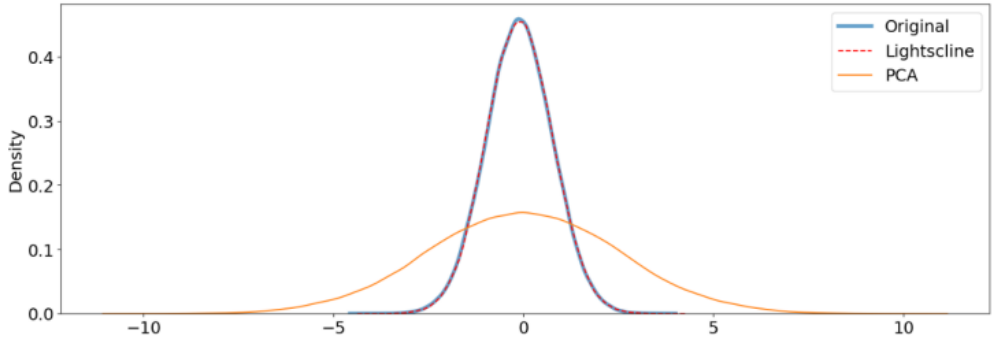
Our Reference
[redacted]

Tel
[redacted]

[redacted]

Version V1.0
Topic **Lightscline**
Description Exploration of Lightscline Free Version

Data reduction technique was a good approach it eliminated redundant information and focused solely on relevant data. The technique was compared with Principal Component Analysis (PCA) and the distributions were found to be different. Lightscline's approach to data reduction kept the distribution of the data intact. Plots below shows the comparison with PCA.



Lightscline
Data Reduction AI

Human Activity Recognition on wearables

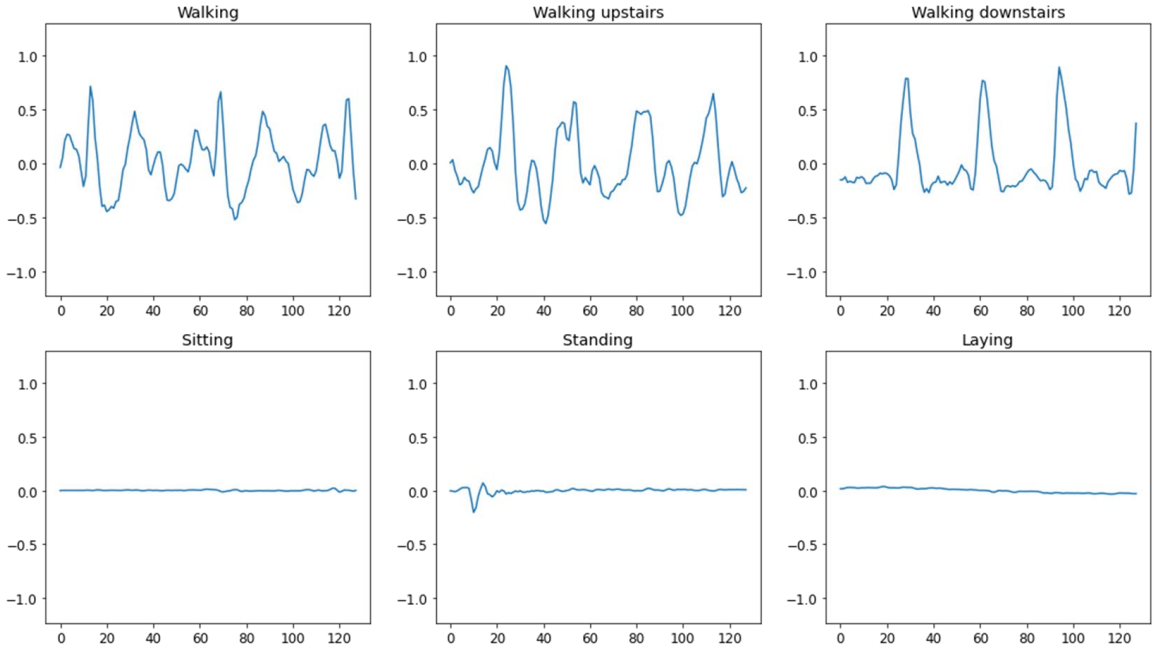
Activities performed by subjects (in seconds)

	subject101	subject102	subject103	subject104	subject105	subject106	subject107	subject108	subject109	Sum	Nr. of subjects
1 – lying	271.86	234.29	220.43	230.46	236.98	233.39	256.1	241.64	0	1925.15	8
2 – sitting	234.79	223.44	287.6	254.91	268.63	230.4	122.81	229.22	0	1851.8	8
3 – standing	217.16	255.75	205.32	247.05	221.31	243.55	257.5	251.59	0	1899.23	8
4 – walking	222.52	325.32	290.35	319.31	320.32	257.2	337.19	315.32	0	2387.53	8
5 – running	212.64	92.37	0	0	246.45	228.24	36.91	165.31	0	981.92	6
6 – cycling	235.74	251.07	0	226.98	245.76	204.85	226.79	254.74	0	1645.93	7
7 – Nordic walking	202.64	297.38	0	275.32	262.7	266.85	287.24	288.87	0	1881	7
9 – watching TV	836.45	0	0	0	0	0	0	0	0	836.45	1
10 – computer work	0	0	0	0	1108.82	617.76	0	687.24	685.49	3099.31	4
11 – car driving	545.18	0	0	0	0	0	0	0	0	545.18	1
12 – ascending stairs	158.88	173.4	103.87	166.92	142.79	132.89	176.44	116.81	0	1172	8
13 – descending stairs	148.97	152.11	152.72	142.83	127.25	112.7	116.16	96.53	0	1049.27	8
16 – vacuum cleaning	229.4	206.82	203.24	200.36	244.44	210.77	215.51	242.91	0	1753.45	8
17 – ironing	235.72	288.79	279.74	249.94	330.33	377.43	294.98	329.89	0	2386.82	8
18 – folding laundry	271.13	0	0	0	0	217.85	0	236.49	273.27	998.74	4
19 – house cleaning	540.88	0	0	0	284.87	287.13	0	416.9	342.05	1871.83	5
20 – playing soccer	0	0	0	0	0	0	0	181.24	287.88	469.12	2
24 – rope jumping	129.11	132.61	0	0	77.32	2.55	0	88.05	63.9	493.54	6
Labeled total	4693.07	2633.35	1743.27	2314.08	4117.97	3623.56	2327.63	4142.75	1652.59	27248.27	
Total	6957.67	4469.99	2528.32	3295.75	5295.54	4917.78	3135.98	5884.41	2019.47	38504.91	

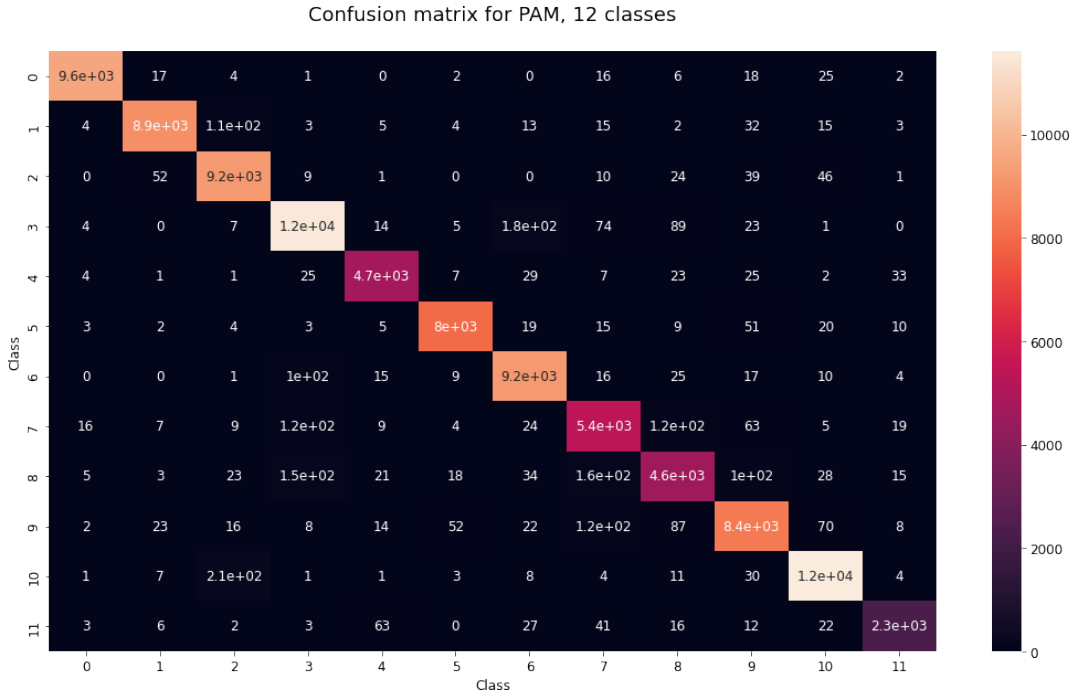
Human Activity Recognition

Model size reduction by ~9x and significant speed-ups

93% accuracy on UCI HAPT dataset (6 classes)



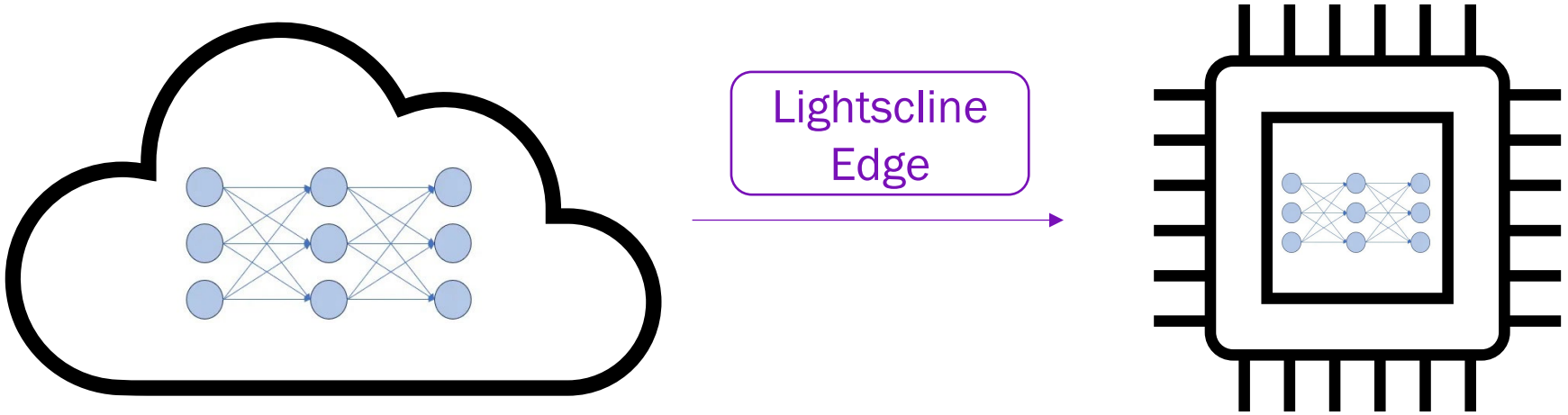
96.7% accuracy on PAMAP2 dataset (12/18 classes)



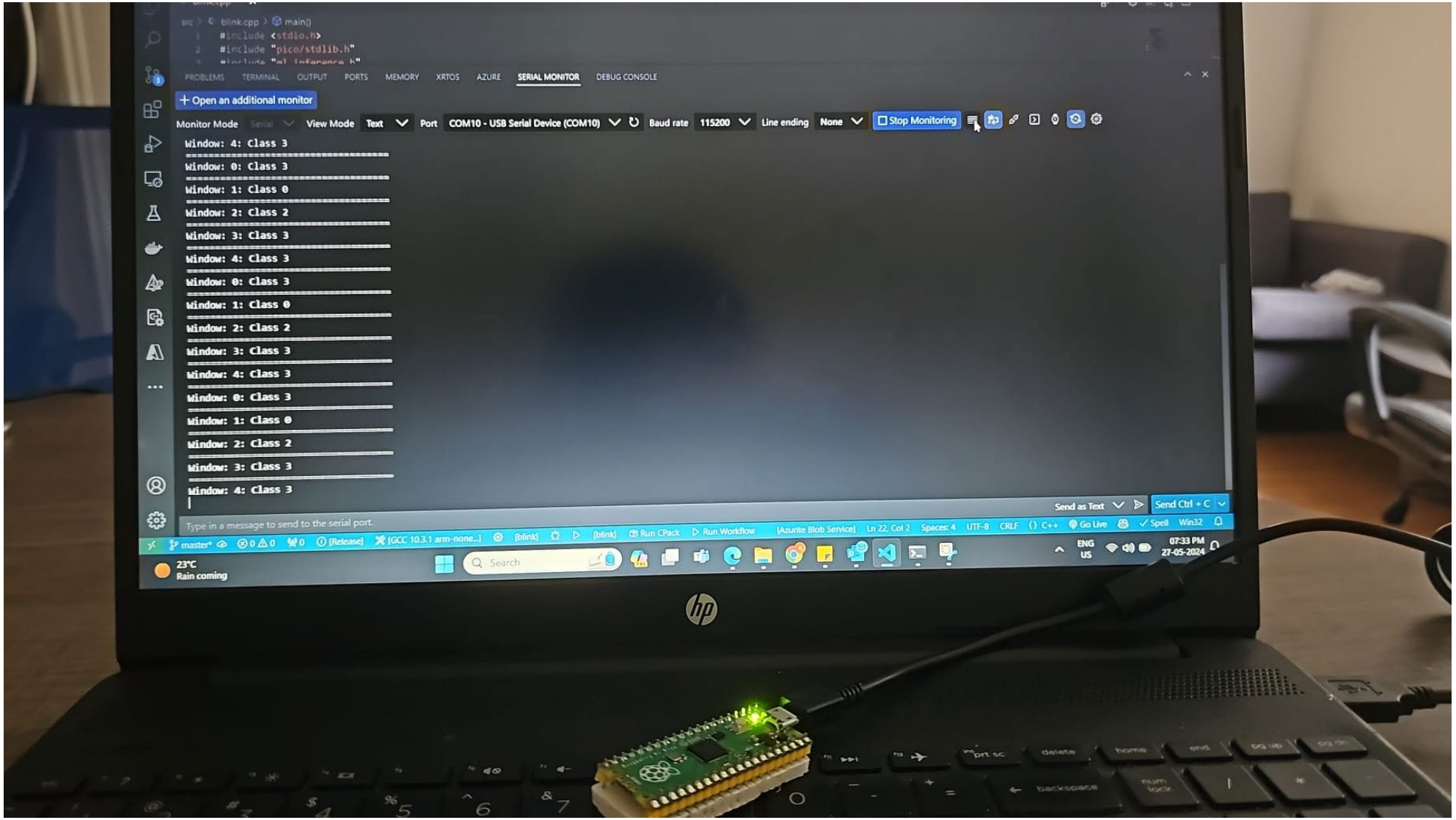
Lightscline
Data Reduction AI

Image cc: https://www.nasa.gov/gallery/hrp-images/

Lightscline Edge

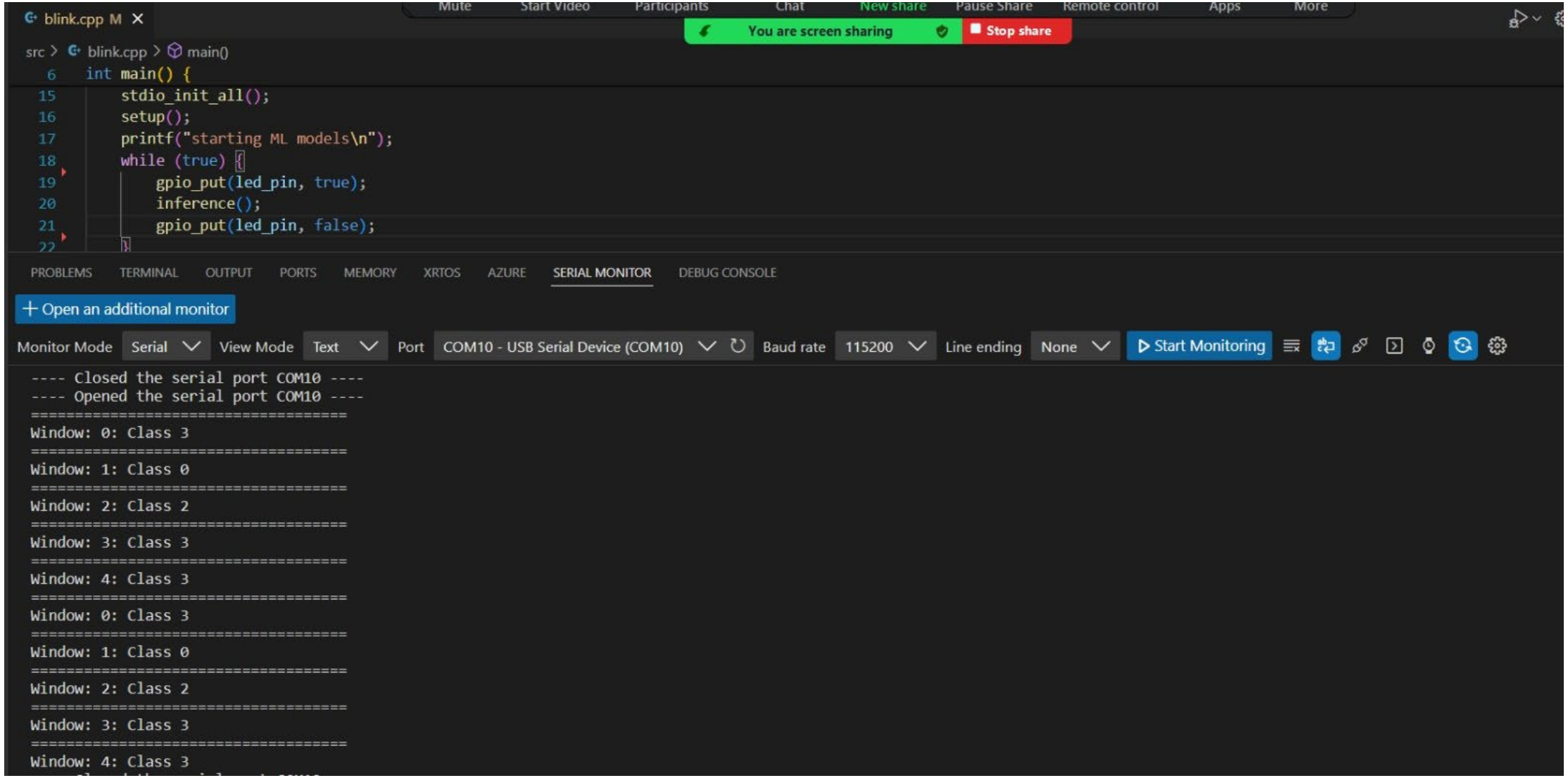


Edge deployment - <256 kB RAM, ~mW power



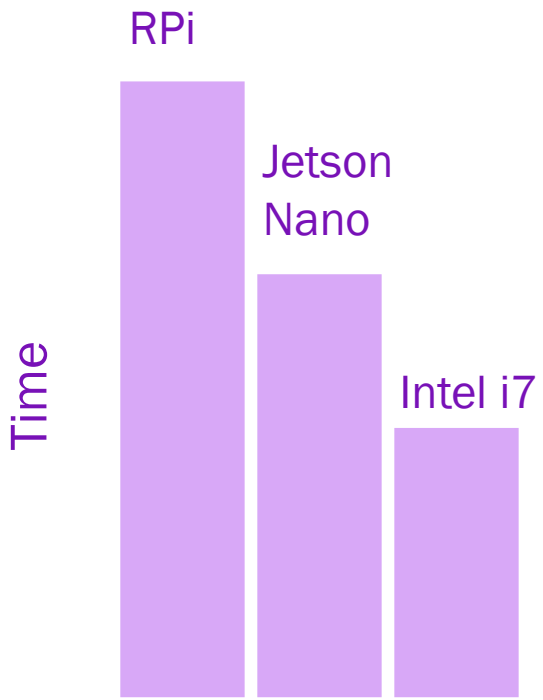
Lightscline
Data Reduction AI

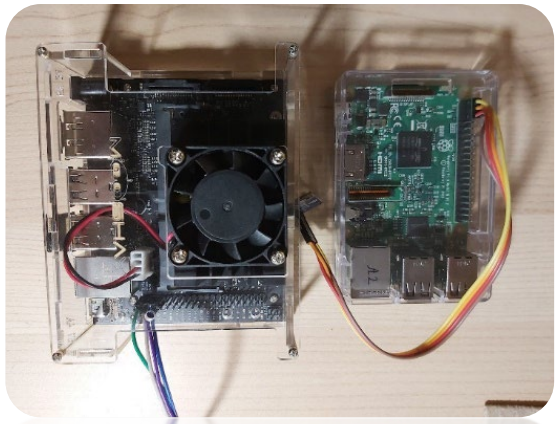
Edge deployment - <256 kB RAM, ~mW power

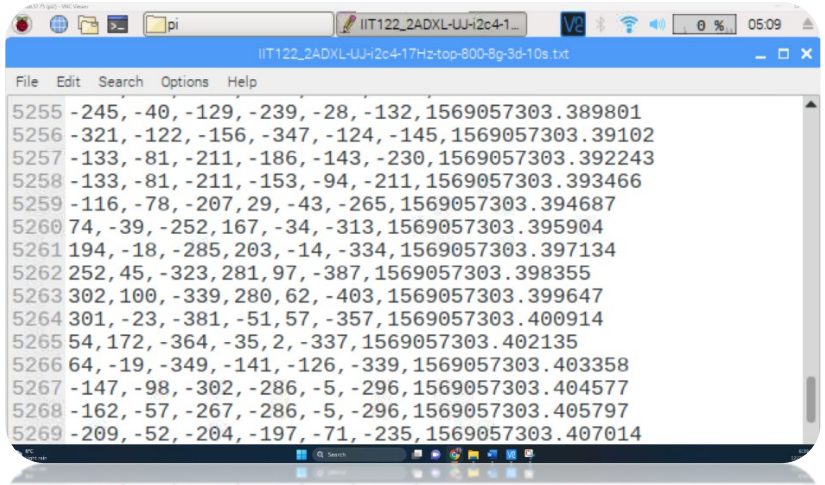


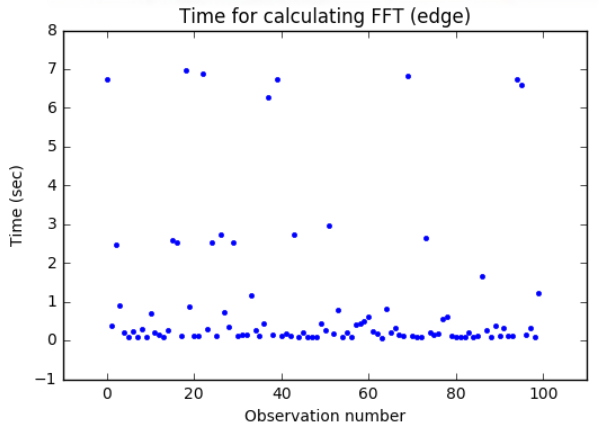
Lightscline
Data Reduction AI

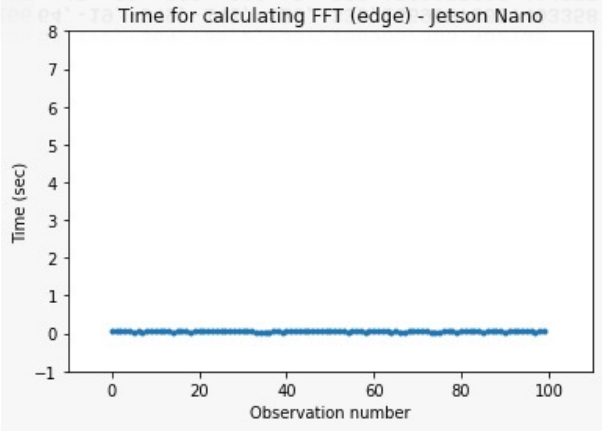
We deploy your models as fast as they are trained!







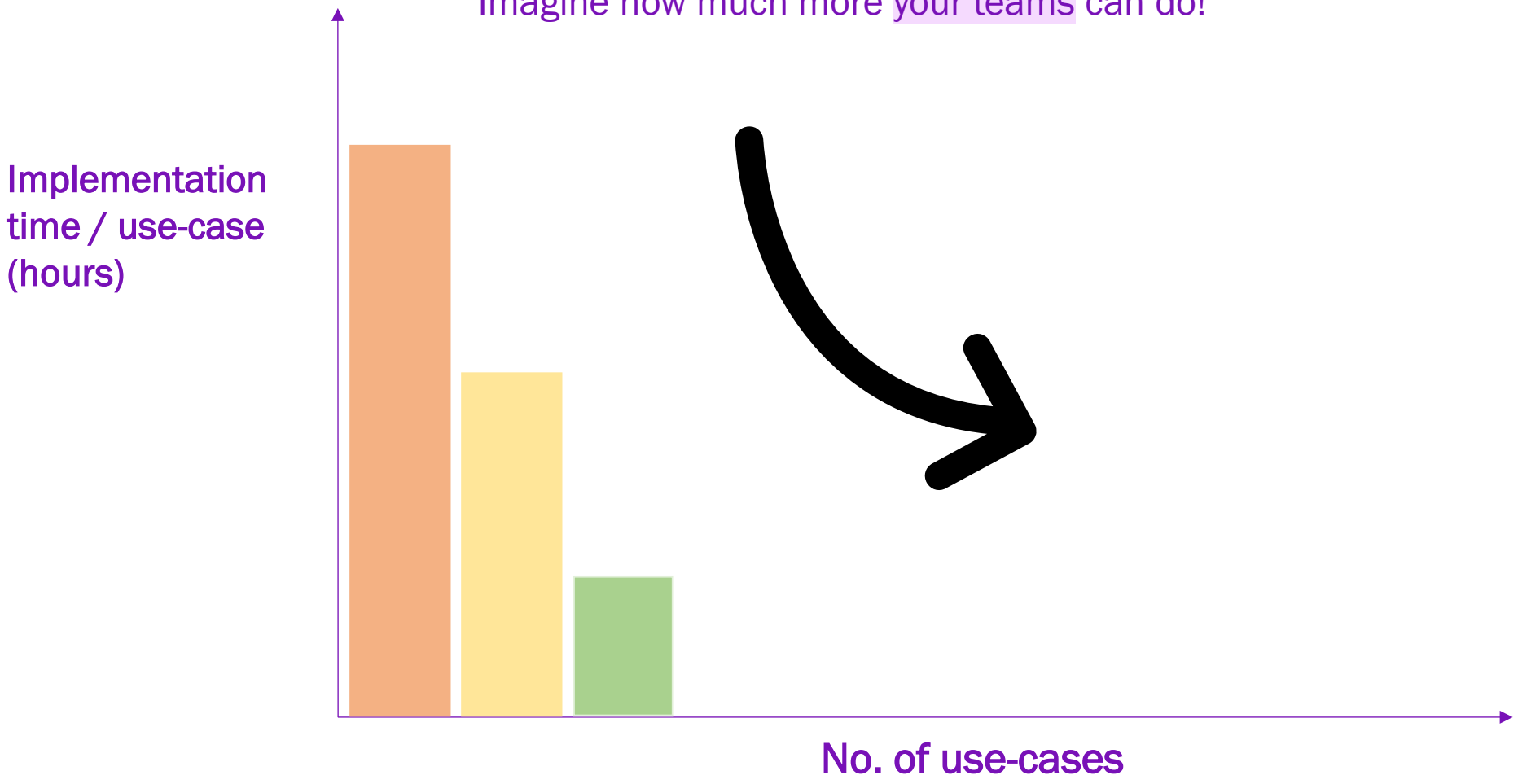




Lightscline
Data Reduction AI

We 10x’ed our productivity by quickly scaling to different use-cases

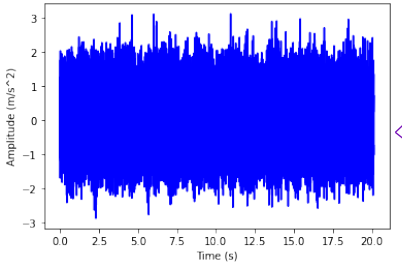
Imagine how much more your teams can do!



Parallel workflow for preserving raw data

Active real-time pipeline: 10x less training & inference costs

Raw data



Lightscline workflow

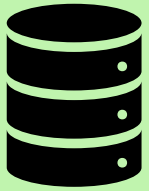
```
from lightscline.lightscline import LightsclineEdge
## Load data into Lightscline
ls = Lightscline(data=data,fs = SAMPLING_FREQUENCY)
## Reduce the amount of data by 70% of the original
ls.reduce_and_preprocess_data(per_reduction=70)
## Train the model
ls.train_model(verbose=True,n_iters = 1000)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

Results

97% accuracy

Cold pipeline: store all raw data

Archival storage

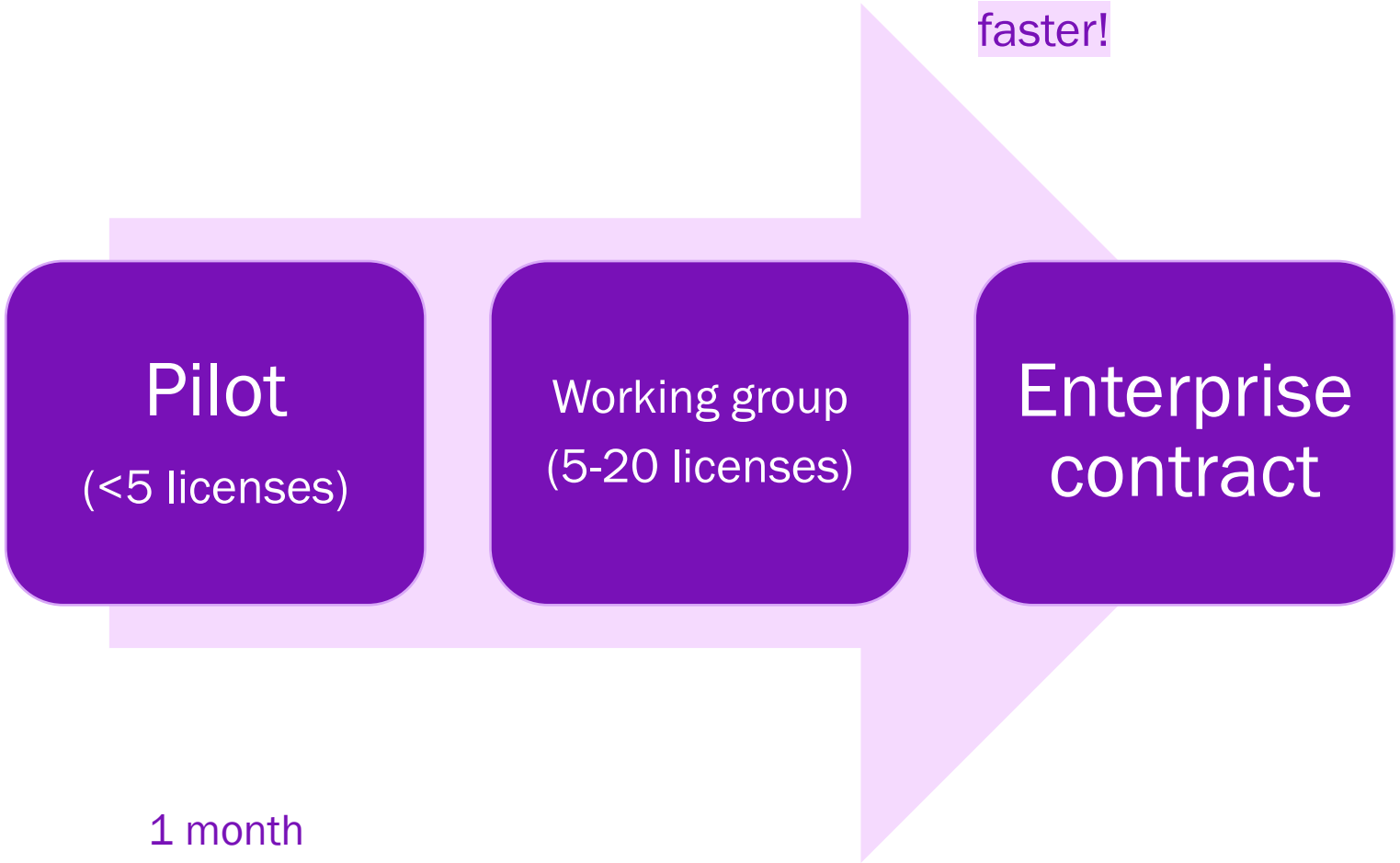


Data is stored for non-real-time analysis

Lightscline
Data Reduction AI

Customer journey

We accelerate your AI journeys by 10 years by helping data teams deploy 10x more, faster!



Lightscline
Data Reduction AI

Engagement with Lightscline

Off the shelf product:

End-to-end predictions

Customized product:

We will work with you to solve your use-case and modify our solutions accordingly

Team

Ankur Verma, PhD

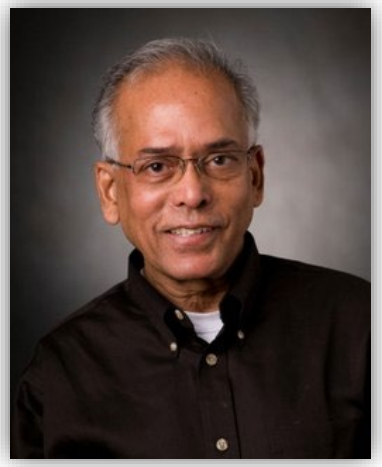
Co-Founder & CEO,

30 under 30 - Society of Mfg Engg,

Ex-Fraunhofer, GM R&D

Ayush Goyal

Co-Founder & CTO, Ex-UHG



Prof. Soundar Kumara

Co-Founder & Chief Scientist,

Allen E. Pearce and Allen M. Pearce

Professor of Industrial Engineering, Father

of Smart Manufacturing