

Usages of Contrastive Losses, Introduction and Applications

for sensor data

Halil Beglerovic & Jörg Simon, June 30th 2020

About me

- PhD on using DeepLearning to detect Human Factors from BioSignals
- Prof. Eduardo Veas and Herbert Danzinger
- Sometimes very Sparse Data!
- Bring fractioned Data Sets together with Contrastive Loss





Deep Learning Troubles

No Labels? Fragmented Datasets?

Nasty Noise Effects?

Time Shifts have huge effect? Clustering?

Contrastive Loss*

*** is partly a solution for some of these problems**

Contrastive Loss*

*** is partly a solution for some of these problems**

we already have seen clustering & time shifts

Contrastive Loss*

*** is partly a solution for some of these problems**

- ~~clustering & time shifts~~
- little labels
- fragmented datasets
- noise

Agenda

- History and other terms for the same
- Fractioned datasets of drowsiness
- Unifying sensor-recordings for later supervised training
 - Demo & Code
- Summary

Agenda

- History and other terms for the same
 - Fractioned datasets of drowsiness
 - Unifying sensor-recordings for later supervised training
 - Demo & Code
 - Summary
- 
- few labels
 - fragmented datasets
 - noise

History and other terms for the same

2005 CVPR: Face Verification

Chopra, S., Hadsell, R., LeCun, Y. (2005). Learning a Similarity Metric Discriminatively, with Application to Face Verification 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) 1, 539-546. <https://dx.doi.org/10.1109/cvpr.2005.202>

History and other terms for the same



Sumit Chopra



Raia Hadsell



Yann LeCun

2005 CVPR: Face Verification

*Chopra, S., Hadsell, R., LeCun, Y. (2005). **Learning a Similarity Metric Discriminatively, with Application to Face Verification** 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) 1, 539-546. <https://dx.doi.org/10.1109/cvpr.2005.202>*

History and other terms for the same



Sumit Chopra



Raia Hadsell



Yann LeCun

2005 CVPR: Face Verification



imposter pair

genuine pair

History and other terms for the same



Sumit Chopra



Raia Hadsell



Yann LeCun

2005 CVPR: Face Verification

$$\mathcal{L}(W) = \sum_{i=1}^P L\left(W, (Y, X_1, X_2)^i\right)$$
$$L\left(W, (Y, X_1, X_2)^i\right) = (1 - Y)L_G\left(E_W(X_1, X_2)^i\right) + YL_I\left(E_W(X_1, X_2)^i\right)$$



imposter pair

genuine pair

History and other terms for the same



Sumit Chopra



Raia Hadsell



Yann LeCun

2005 CVPR: Face Verification

Contrastive Loss

$$\mathcal{L}(W) = \sum_{i=1}^P L\left(W, (Y, X_1, X_2)^i\right)$$
$$L\left(W, (Y, X_1, X_2)^i\right) = (1 - Y)L_G\left(E_W(X_1, X_2)^i\right) + YL_I\left(E_W(X_1, X_2)^i\right)$$



imposter pair

genuine pair

Chopra, S., Hadsell, R., LeCun, Y. (2005). **Learning a Similarity Metric Discriminatively, with Application to Face Verification** 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) 1, 539-546. <https://dx.doi.org/10.1109/cvpr.2005.202>

History and other terms for the same



Sumit Chopra



Raia Hadsell



Yann LeCun

2005 CVPR: Face Verification

$$\mathcal{L}(W) = \sum_{i=1}^P L(W, (Y, X_1, X_2)^i)$$

$$L(W, (Y, X_1, X_2)^i) = (1 - Y)L_G(E_W(X_1, X_2)^i) + YL_I(E_W(X_1, X_2)^i)$$

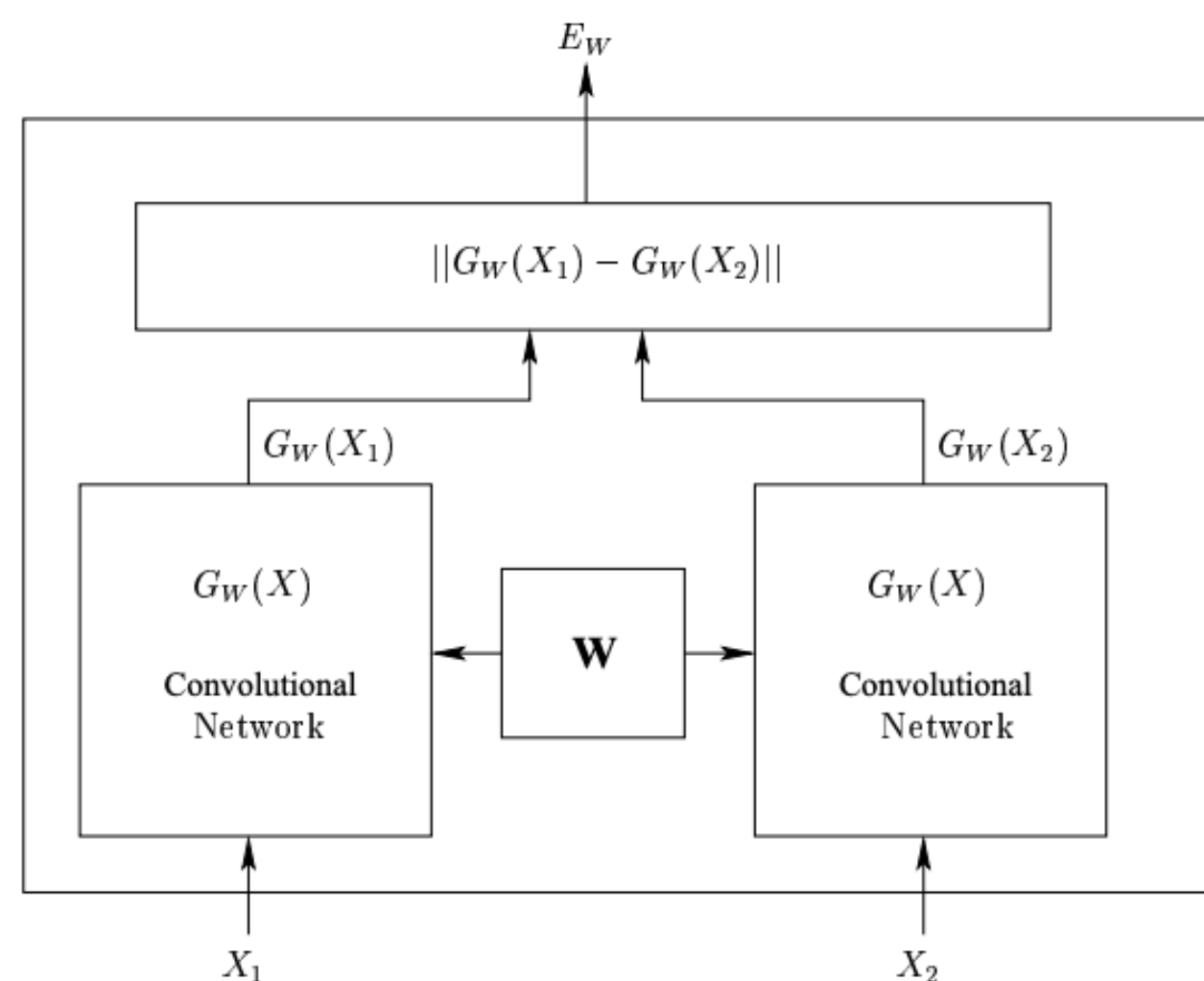


imposter pair

genuine pair

History and other terms for the same

Siamese Networks



$$\mathcal{L}(W) = \sum_{i=1}^P L\left(W, (Y, X_1, X_2)^i\right)$$
$$L\left(W, (Y, X_1, X_2)^i\right) = (1 - Y)L_G\left(E_W(X_1, X_2)^i\right) + YL_I\left(E_W(X_1, X_2)^i\right)$$



Sumit Chopra



Raia Hadsell



Yann LeCun

2005 CVPR: Face Verification



imposter pair

genuine pair

History and other terms for the same

2012 NeurIPS:
Distances on
Images

*Norouzi, M., Fleet, D., Salakhutdinov, R. (2012). **Hamming Distance Metric Learning** 2012 Neural Information Processing Conference (NIPS), 1061–1069, <http://papers.nips.cc/paper/4808-hamming-distance-metric-learning.pdf>*

History and other terms for the same

2012 NeurIPS:
Distances on
Images



Mohammad
Norouzi



David J. Fleet



Ruslan
Salakhutdinov

*Norouzi, M., Fleet, D., Salakhutdinov, R. (2012). **Hamming Distance Metric Learning** 2012 Neural Information Processing Conference (NIPS), 1061–1069, <http://papers.nips.cc/paper/4808-hamming-distance-metric-learning.pdf>*

History and other terms for the same

2012 NeurIPS:
Distances on
Images



Mohammad
Norouzi



David J. Fleet



Ruslan
Salakhutdinov

Triplet Loss

$$\ell_{\text{triplet}}(\mathbf{h}, \mathbf{h}^+, \mathbf{h}^-) = \left[\|\mathbf{h} - \mathbf{h}^+\|_H - \|\mathbf{h} - \mathbf{h}^-\|_H + 1 \right]_+$$

History and other terms for the same



Mohammad
Norouzi

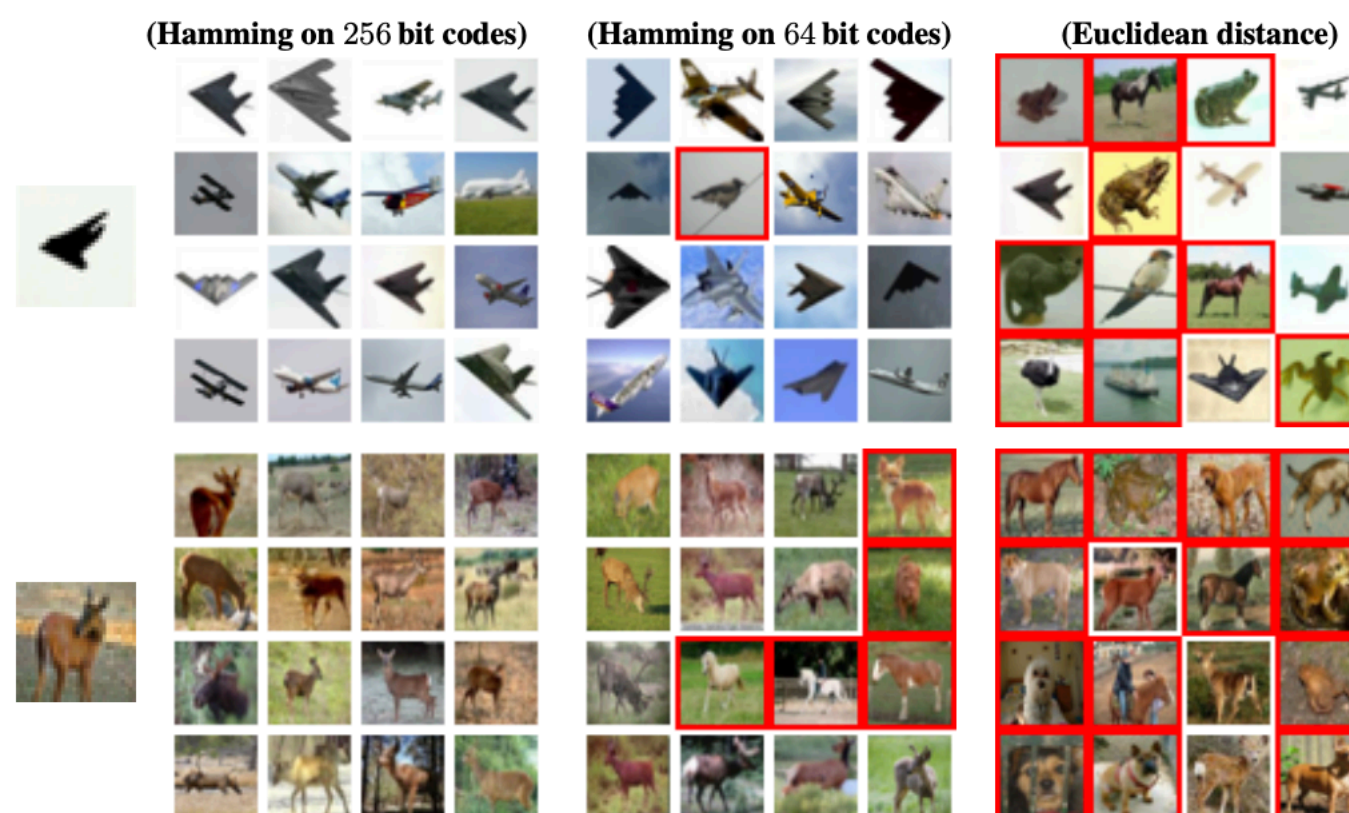


David J. Fleet



Ruslan
Salakhutdinov

2012 NeurIPS: Distances on Images



Triplet Loss

$$\ell_{\text{triplet}}(\mathbf{h}, \mathbf{h}^+, \mathbf{h}^-) = \left[\|\mathbf{h} - \mathbf{h}^+\|_H - \|\mathbf{h} - \mathbf{h}^-\|_H + 1 \right]_+$$

History and other terms for the same

Metric Learning



Mohammad
Norouzi

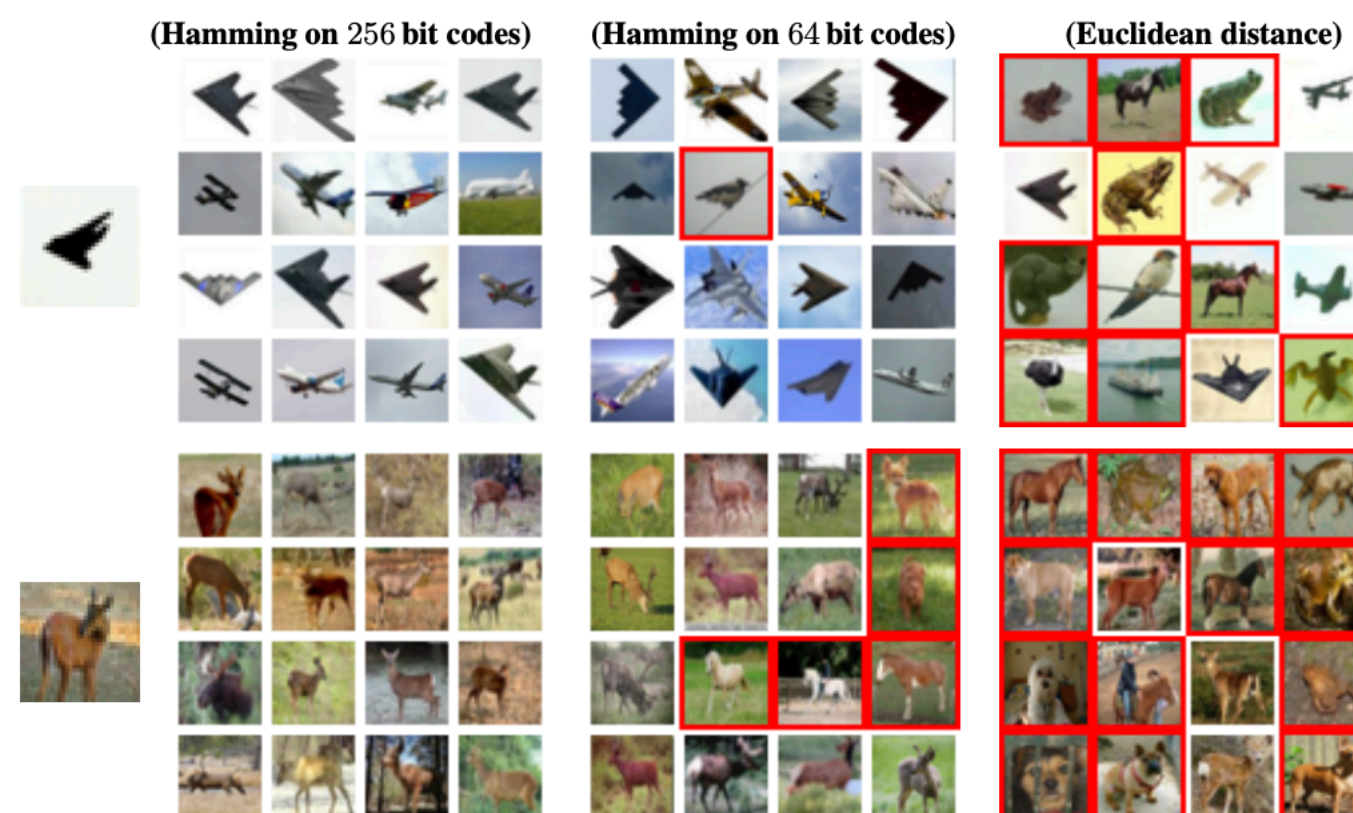


David J. Fleet



Ruslan
Salakhutdinov

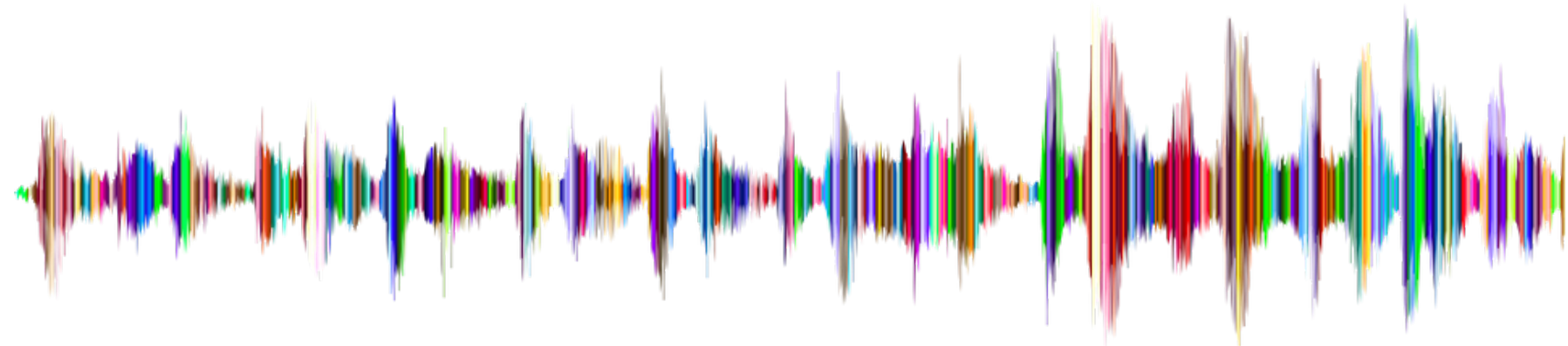
2012 NeurIPS: Distances on Images



Triplet Loss

$$\ell_{\text{triplet}}(\mathbf{h}, \mathbf{h}^+, \mathbf{h}^-) = \left[\|\mathbf{h} - \mathbf{h}^+\|_H - \|\mathbf{h} - \mathbf{h}^-\|_H + 1 \right]_+$$

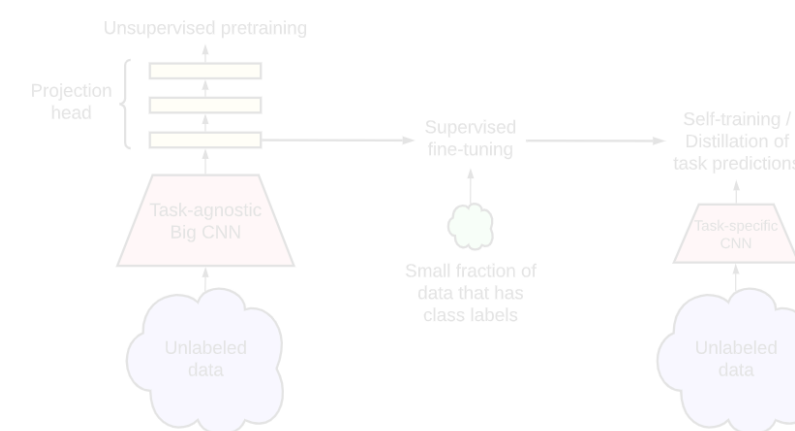
History and other terms for the same



Lian, Z., Li, Y., Tao, J., Huang, J. (2018). Speech Emotion Recognition via Contrastive Loss under Siamese Networks <https://dx.doi.org/10.1145/3267935.3267946>

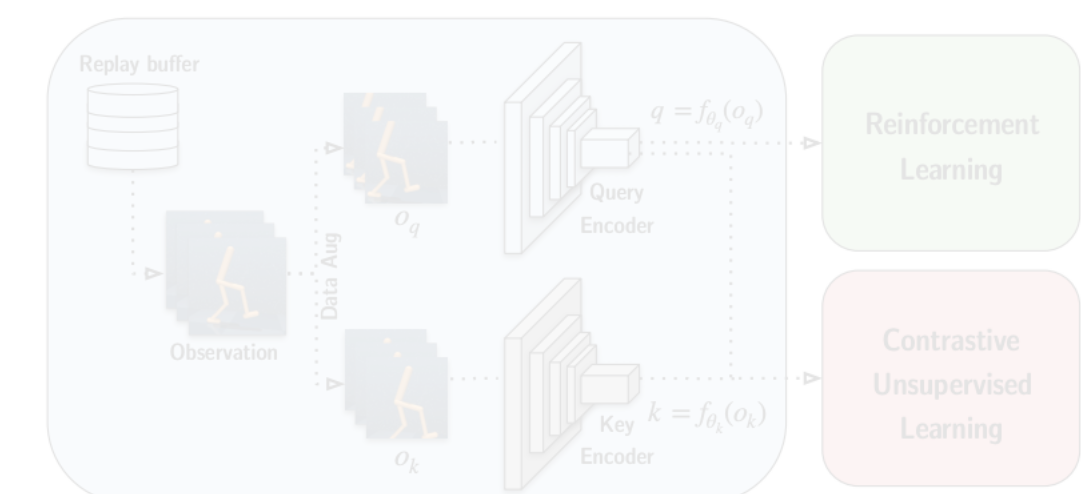
2018 Application to Sensor Data

Chen, T., Kornblith, S., Norouzi, M., Hinton, G. (2020). A Simple Framework for Contrastive Learning of Visual Representations <https://arxiv.org/abs/2002.05709>

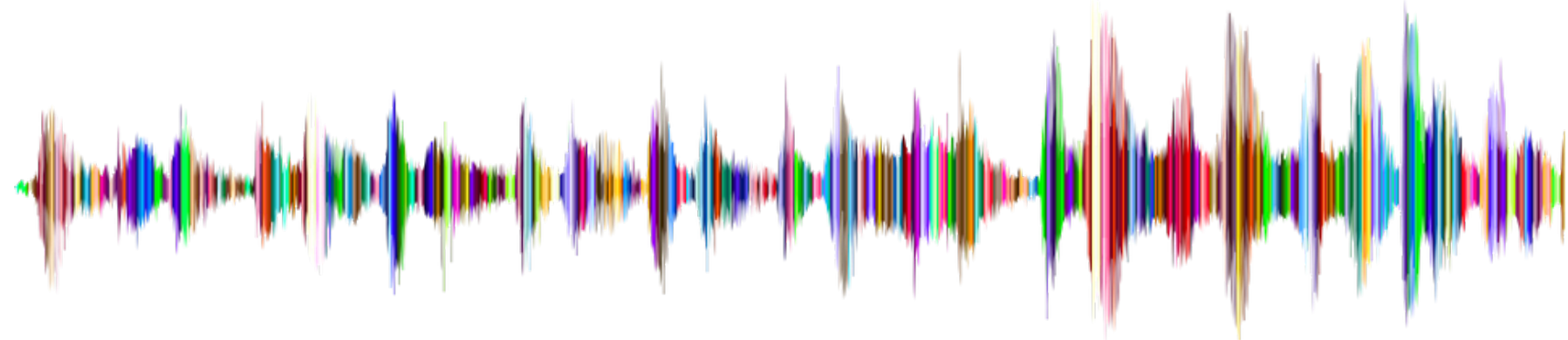


Chen, T., Kornblith, S., Swersky, K., Norouzi, M., Hinton, G. (2020). Big Self-Supervised Models are Strong Semi-Supervised Learners

Srinivas, A., Laskin, M., Abbeel, P. (2020). CURL: Contrastive Unsupervised Representations for Reinforcement Learning arXiv <https://arxiv.org/abs/2004.04136>

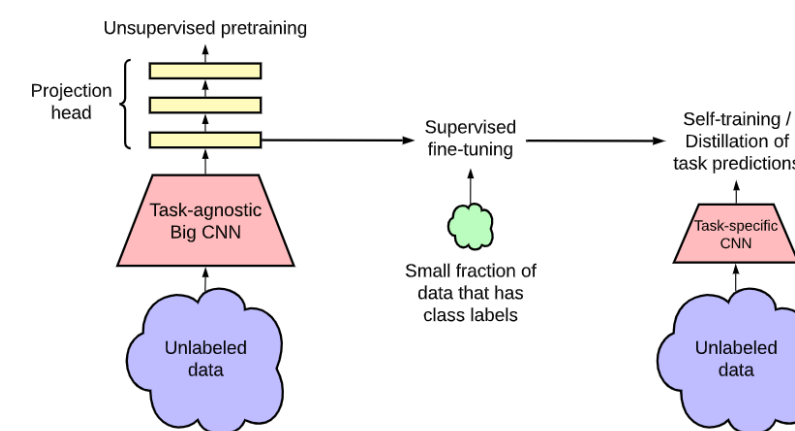


History and other terms for the same



Lian, Z., Li, Y., Tao, J., Huang, J. (2018). Speech Emotion Recognition via Contrastive Loss under Siamese Networks<https://dx.doi.org/10.1145/3267935.3267946>

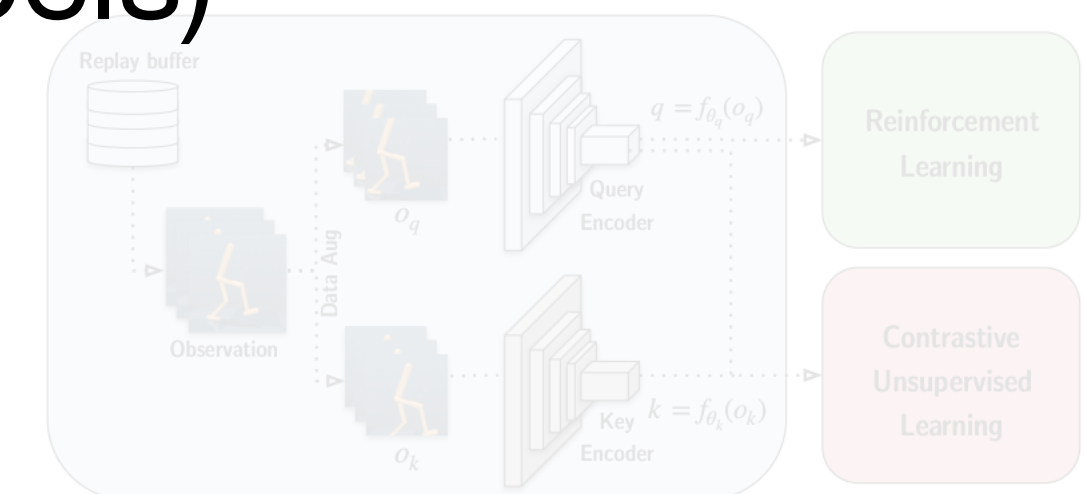
Chen, T., Kornblith, S., Norouzi, M., Hinton, G. (2020). A Simple Framework for Contrastive Learning of Visual Representations<https://arxiv.org/abs/2002.05709>



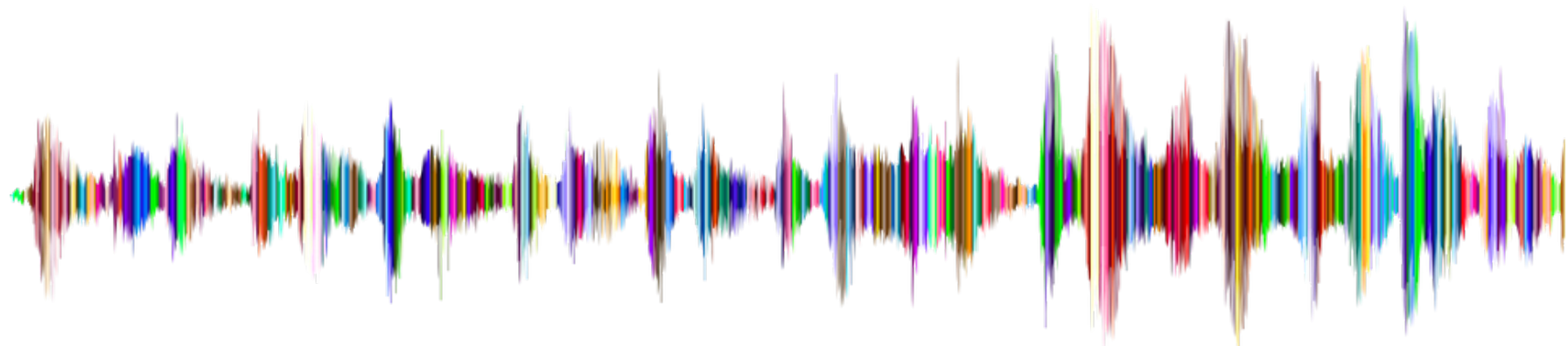
Chen, T., Kornblith, S., Swersky, K., Norouzi, M., Hinton, G. (2020). Big Self-Supervised Models are Strong Semi-Supervised Learners

Early 2020 Sample Efficiency (few labels)

Srinivas, A., Laskin, M., Abbeel, P. (2020). CURL: Contrastive Unsupervised Representations for Reinforcement Learning arXiv <https://arxiv.org/abs/2004.04136>

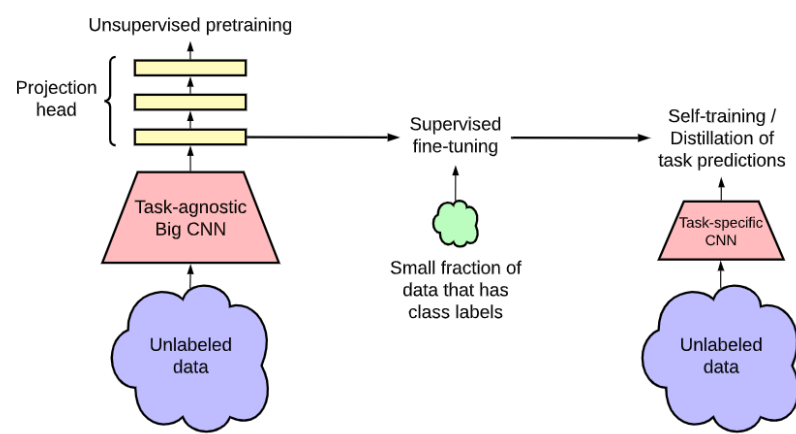


History and other terms for the same



Lian, Z., Li, Y., Tao, J., Huang, J. (2018). **Speech Emotion Recognition via Contrastive Loss under Siamese Networks**<https://dx.doi.org/10.1145/3267935.3267946>

Chen, T., Kornblith, S., Norouzi, M., Hinton, G. (2020). **A Simple Framework for Contrastive Learning of Visual Representations**<https://arxiv.org/abs/2002.05709>

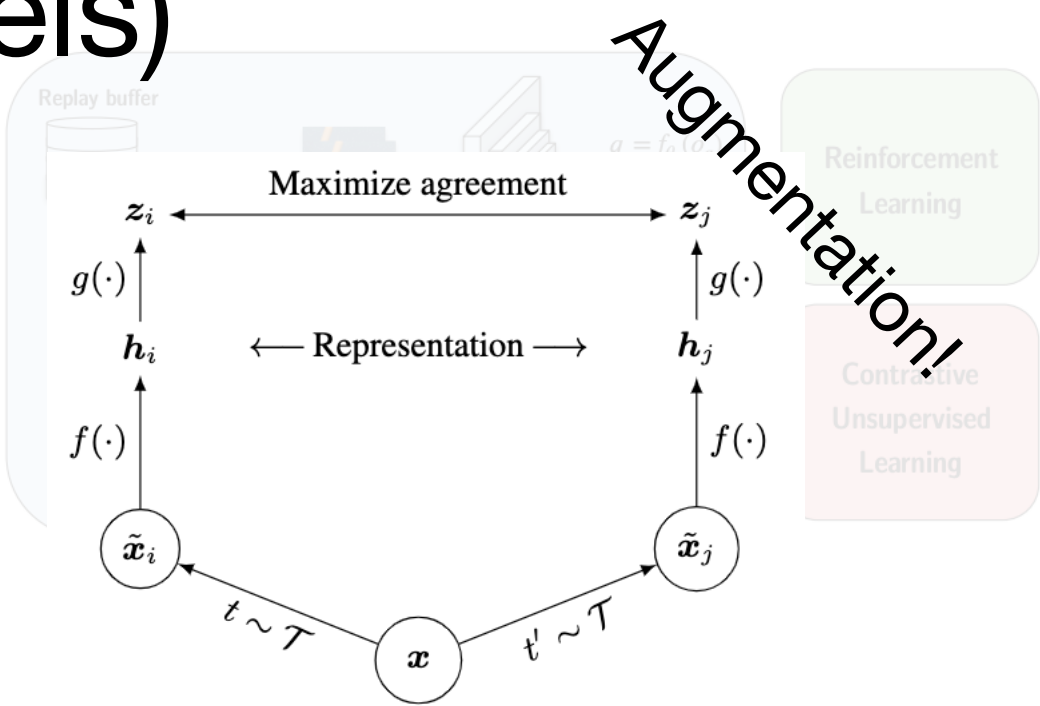


Chen, T., Kornblith, S., Swersky, K., Norouzi, M., Hinton, G. (2020). **Big Self-Supervised Models are Strong Semi-Supervised Learners**

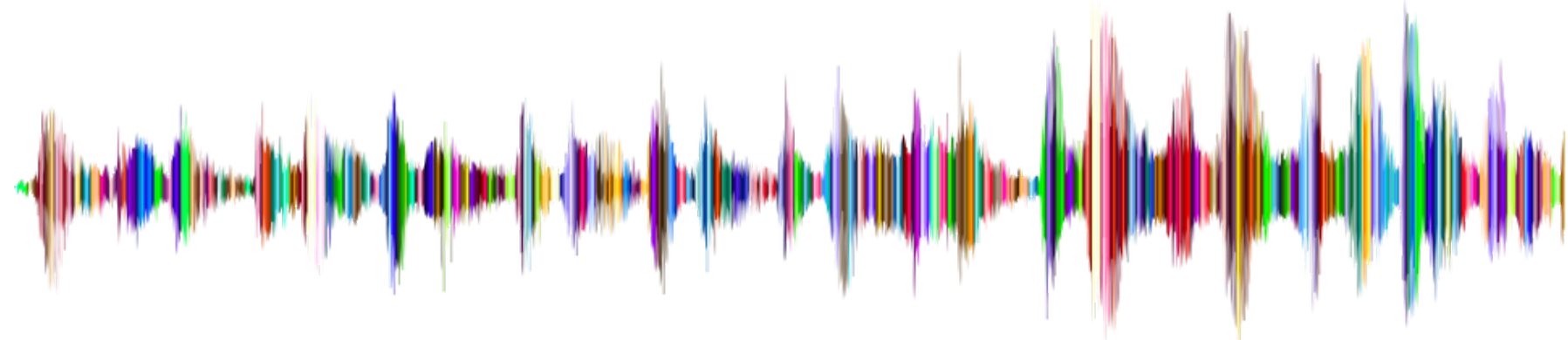
Early 2020 Sample Efficiency (few labels)

Srinivas, A., Laskin, M., Abbeel, P. (2020). **CURL: Contrastive Unsupervised Representations for Reinforcement Learning** arXiv <https://arxiv.org/abs/2004.04136>

$$\ell_{i,j} = -\log \frac{\exp(\text{sim}(\mathbf{z}_i, \mathbf{z}_j)/\tau)}{\sum_{k=1}^{2N} \mathbf{1}_{[k \neq i]} \exp(\text{sim}(\mathbf{z}_i, \mathbf{z}_k)/\tau)}$$

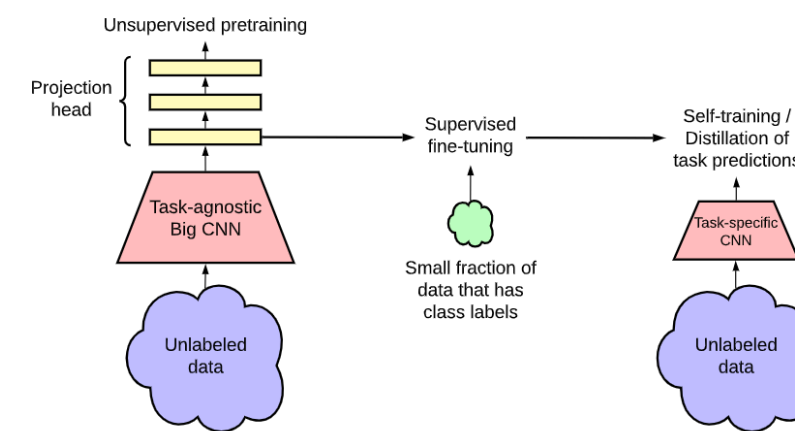


History and other terms for the same



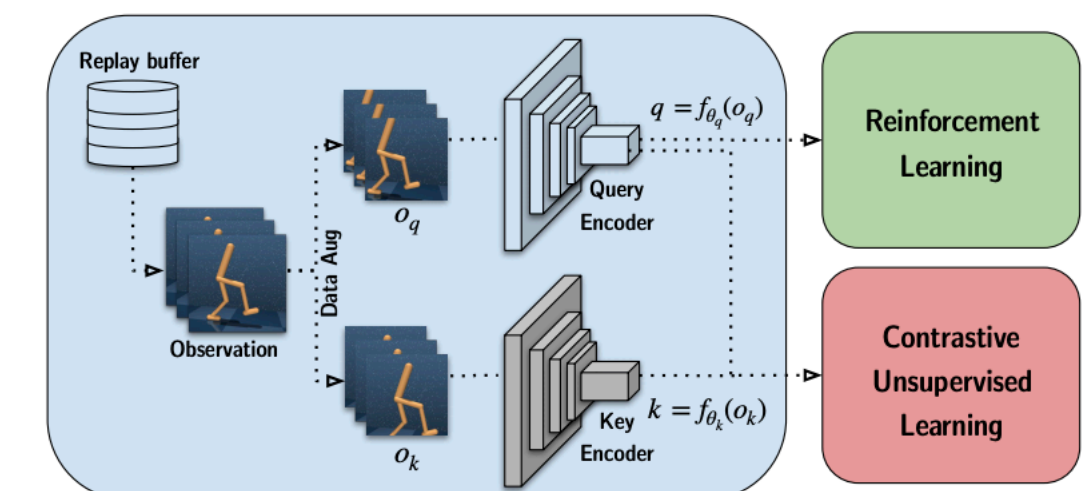
Lian, Z., Li, Y., Tao, J., Huang, J. (2018). **Speech Emotion Recognition via Contrastive Loss under Siamese Networks**<https://dx.doi.org/10.1145/3267935.3267946>

Chen, T., Kornblith, S., Norouzi, M., Hinton, G. (2020). **A Simple Framework for Contrastive Learning of Visual Representations**<https://arxiv.org/abs/2002.05709>



Chen, T., Kornblith, S., Swersky, K., Norouzi, M., Hinton, G. (2020). **Big Self-Supervised Models are Strong Semi-Supervised Learners**

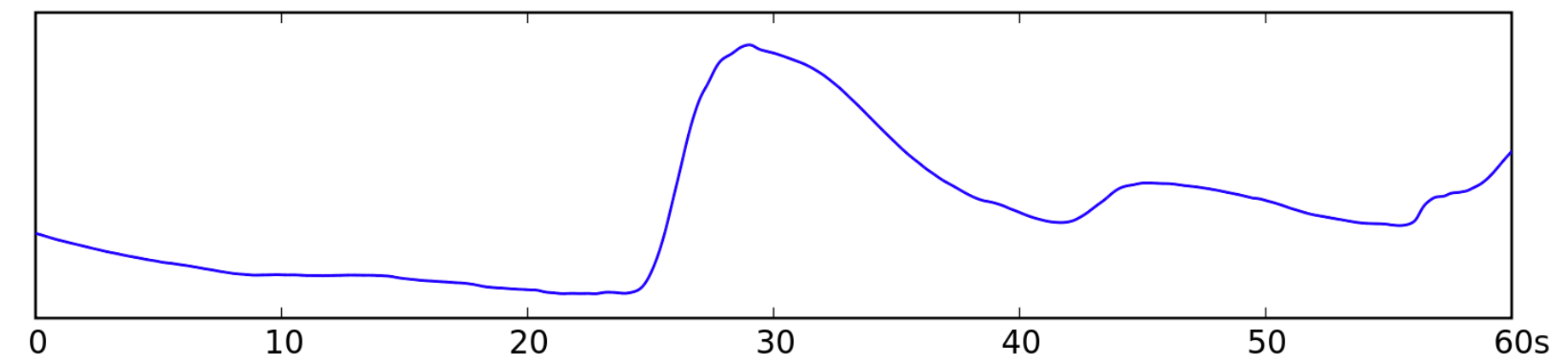
Srinivas, A., Laskin, M., Abbeel, P. (2020). **CURL: Contrastive Unsupervised Representations for Reinforcement Learning** arXiv <https://arxiv.org/abs/2004.04136>



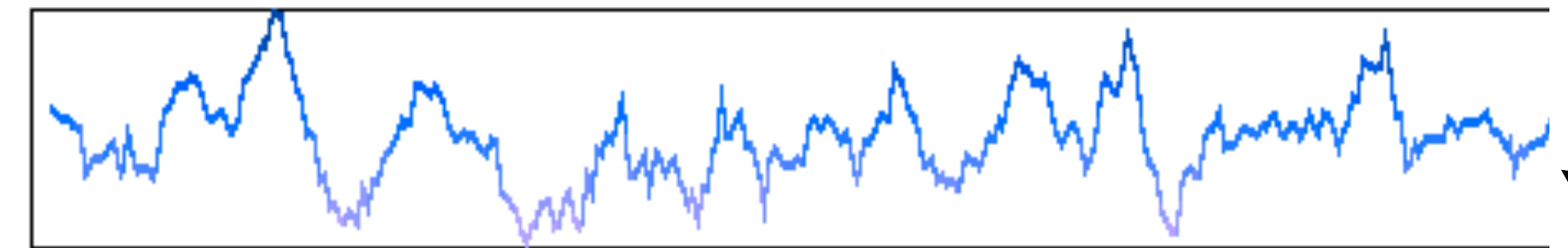
2020 Sample Efficiency in Reinforcement Learning

Fractioned Datasets of Drowsiness

- We have a unit generating a signal
- Unit: f.e. a **driver** of a car
- Signal: **Physiological signals** for modes (alert, normal, drowsy, asleep)
- We have data from 4 experiments, but they all used **slightly different sensors** of the same type (EEG, ECG, GSR)



GSR Signal Sensor 1

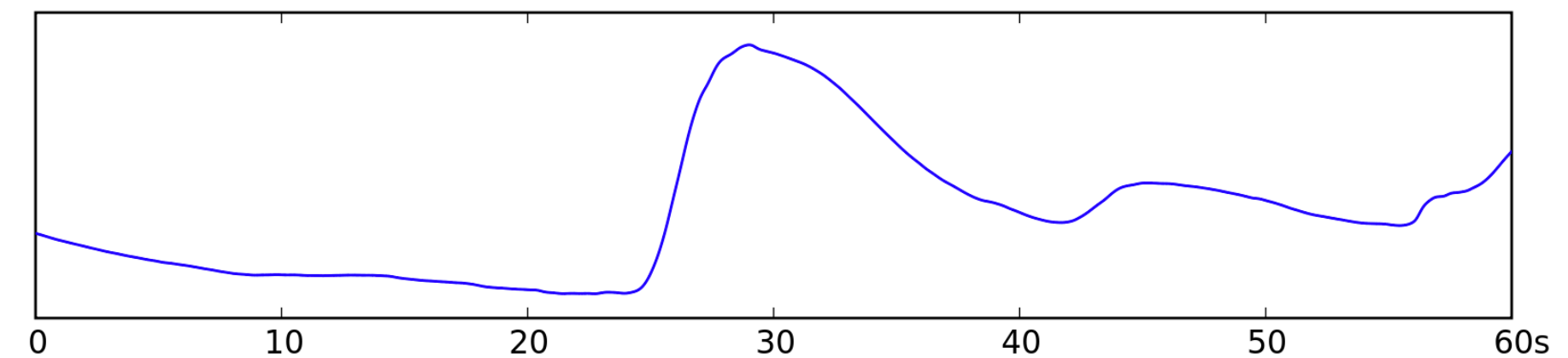


GSR Signal Sensor 2

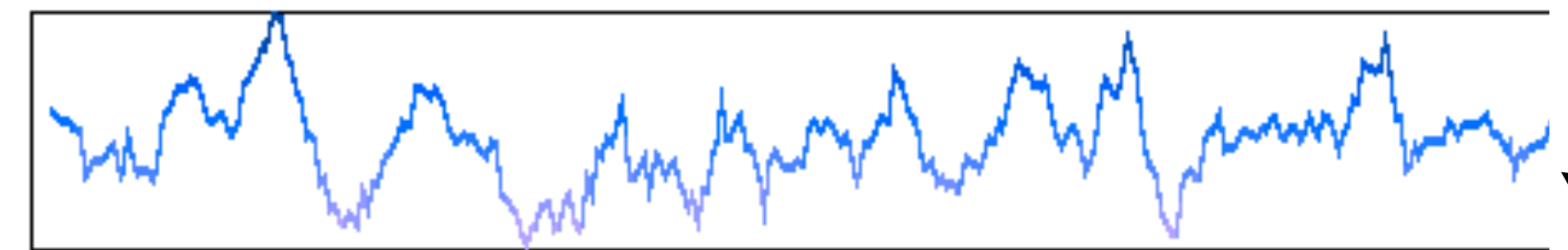
Much more noise!

Fractioned Datasets of Drowsiness

- Contrastive Loss:
- Use all data
- Transform data into embedding where differences of sensors are mitigated
- Fine-tune classifier on that



GSR Signal Sensor 1

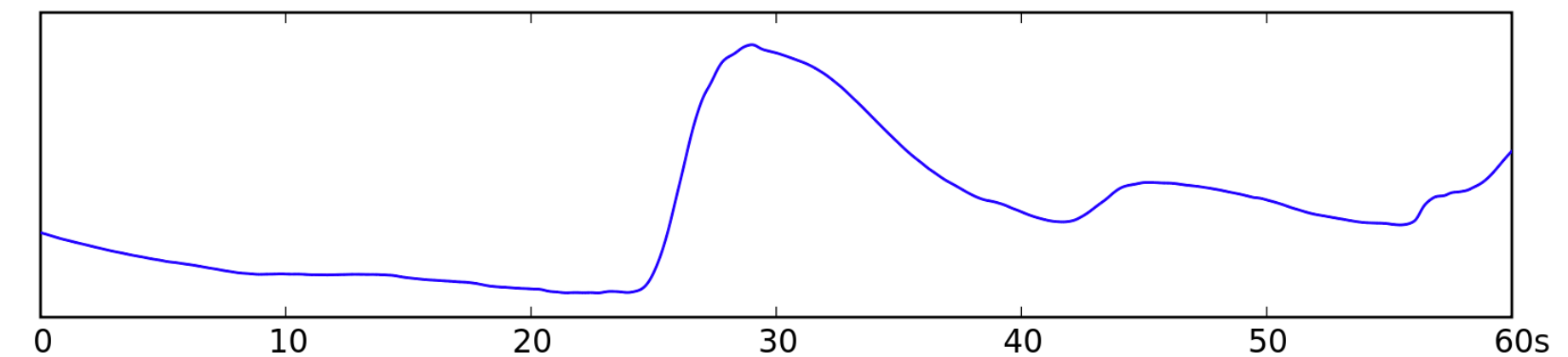


GSR Signal Sensor 2

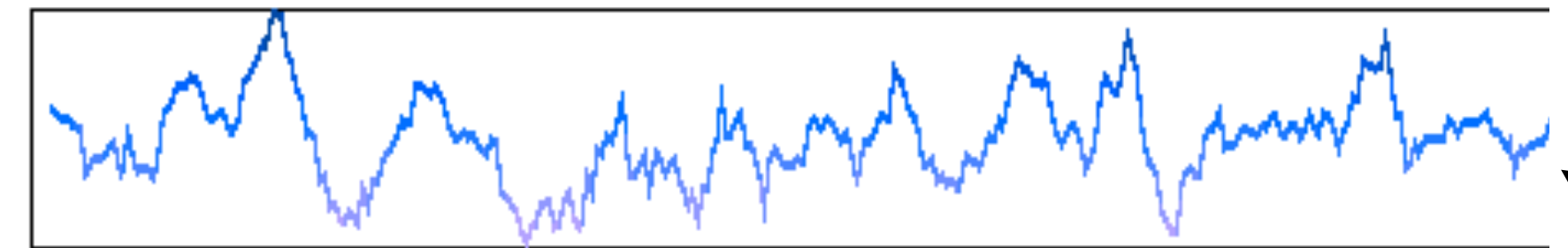
Much more noise!

Fractioned Datasets of Drowsiness

- Bad news: I can not show you exactly this use case (confidential datasets)
- Good news: I have a “surrogate problem”



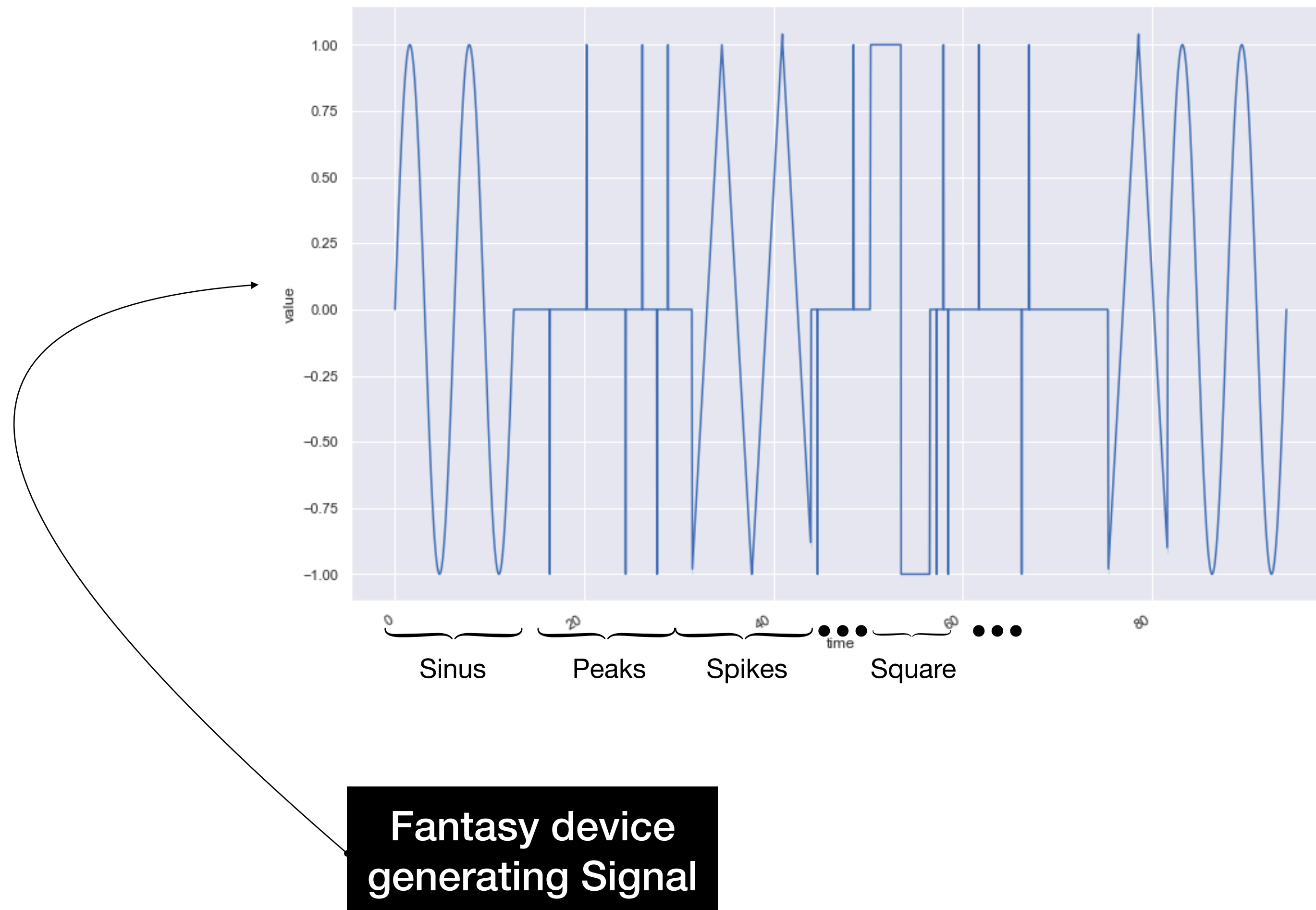
GSR Signal Sensor 1



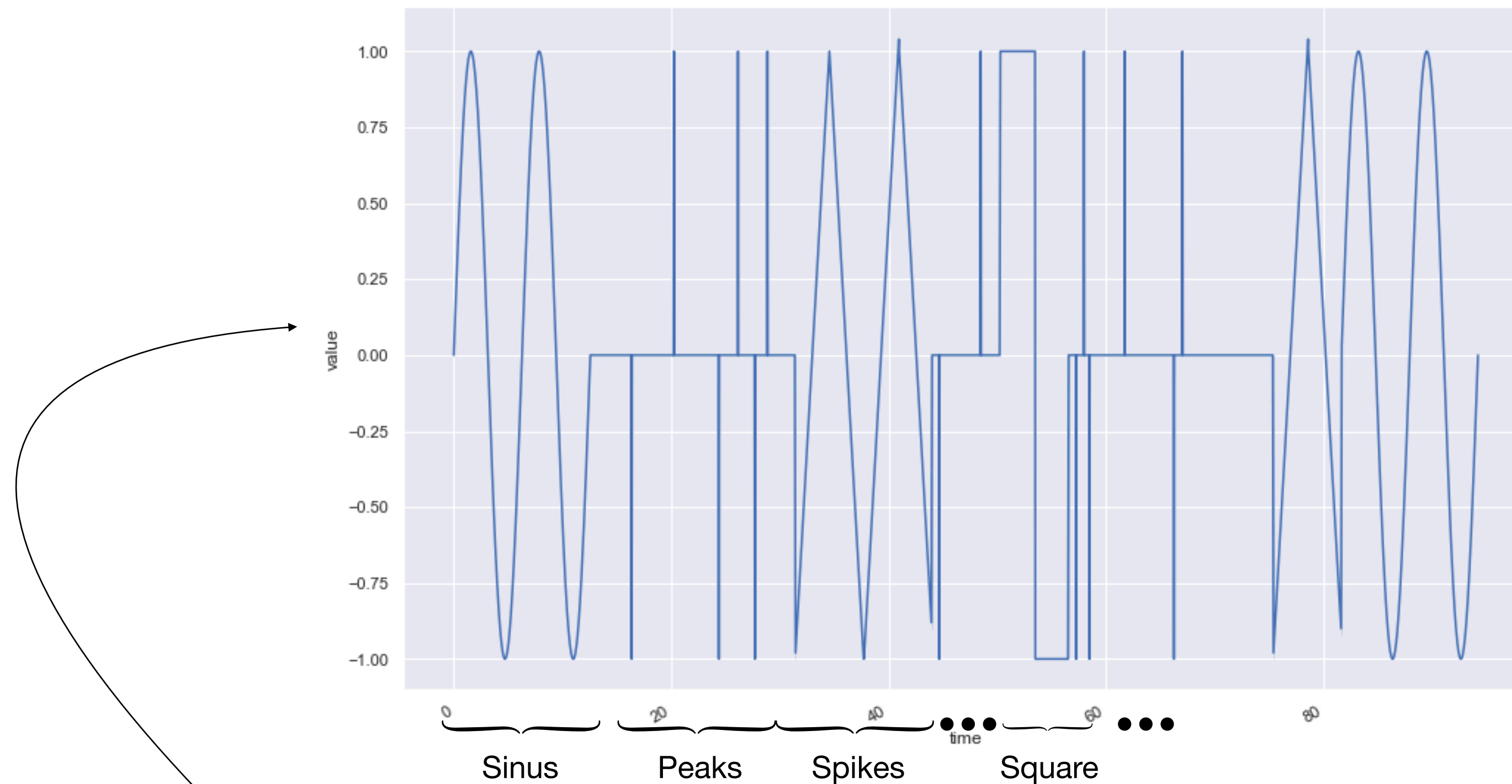
GSR Signal Sensor 2

Much more noise!

Unifying Sensor-Recordings for later Supervised Training



Unifying Sensor-Recordings for later Supervised Training

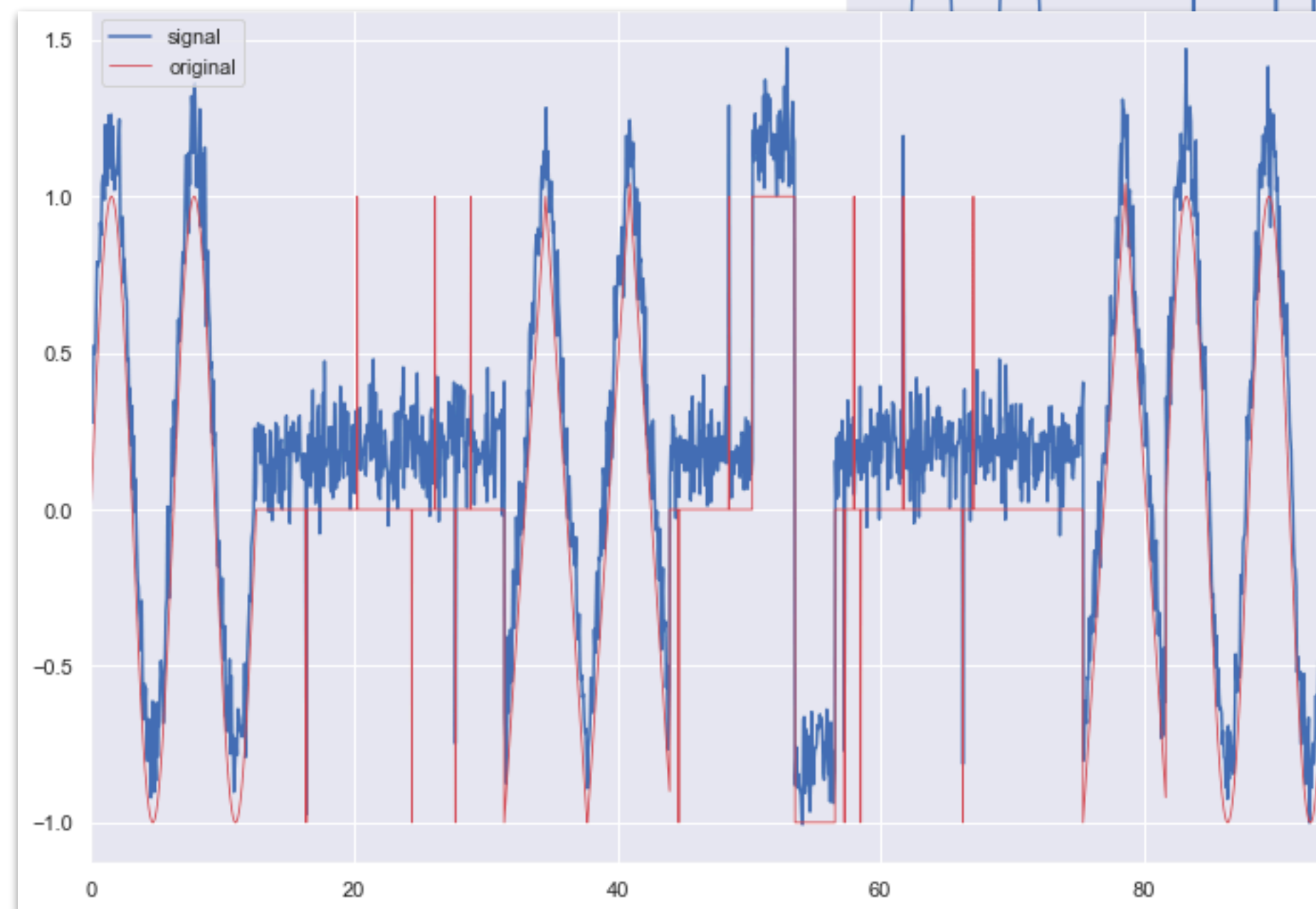


Fantasy device
generating Signal

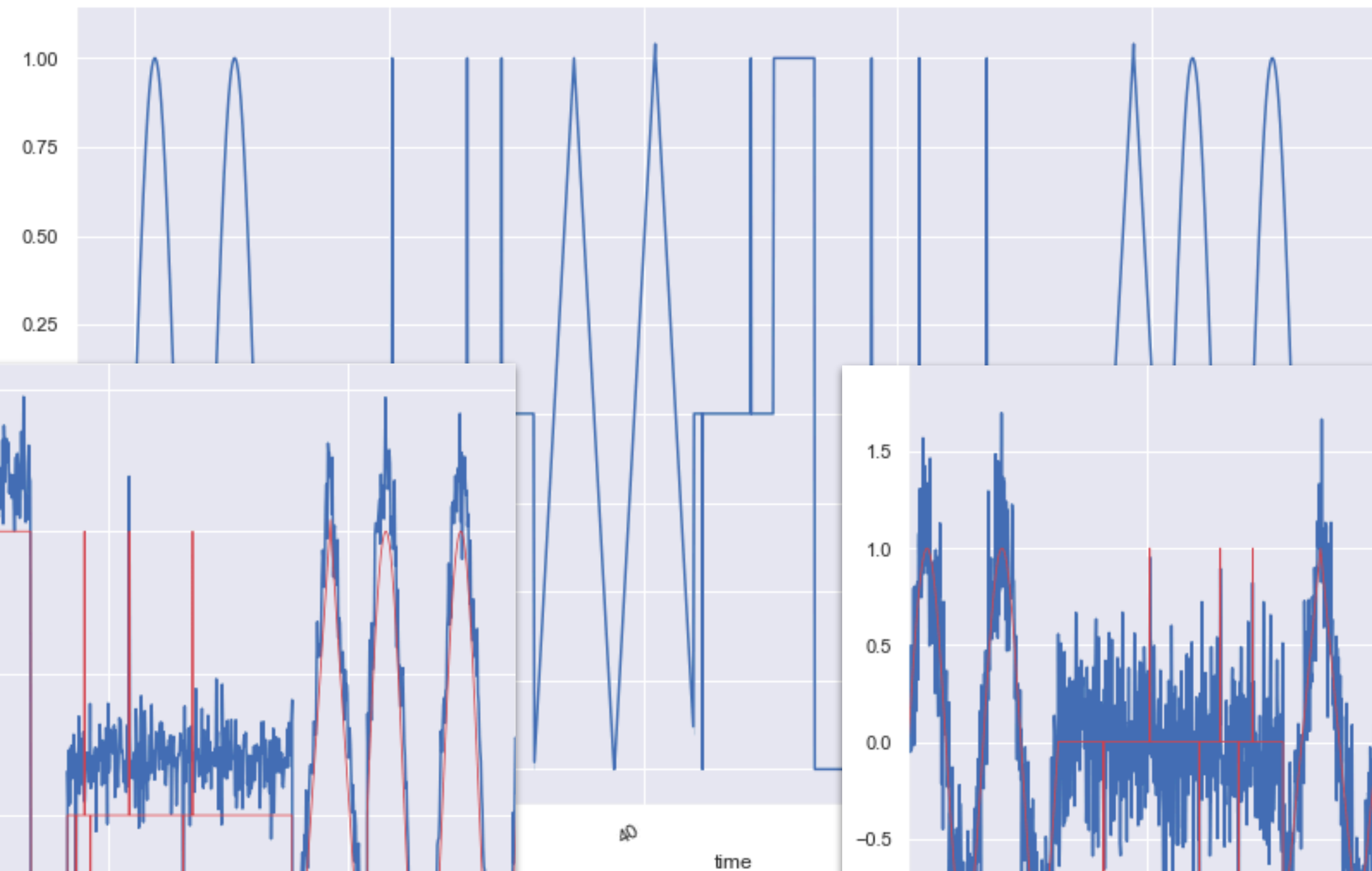
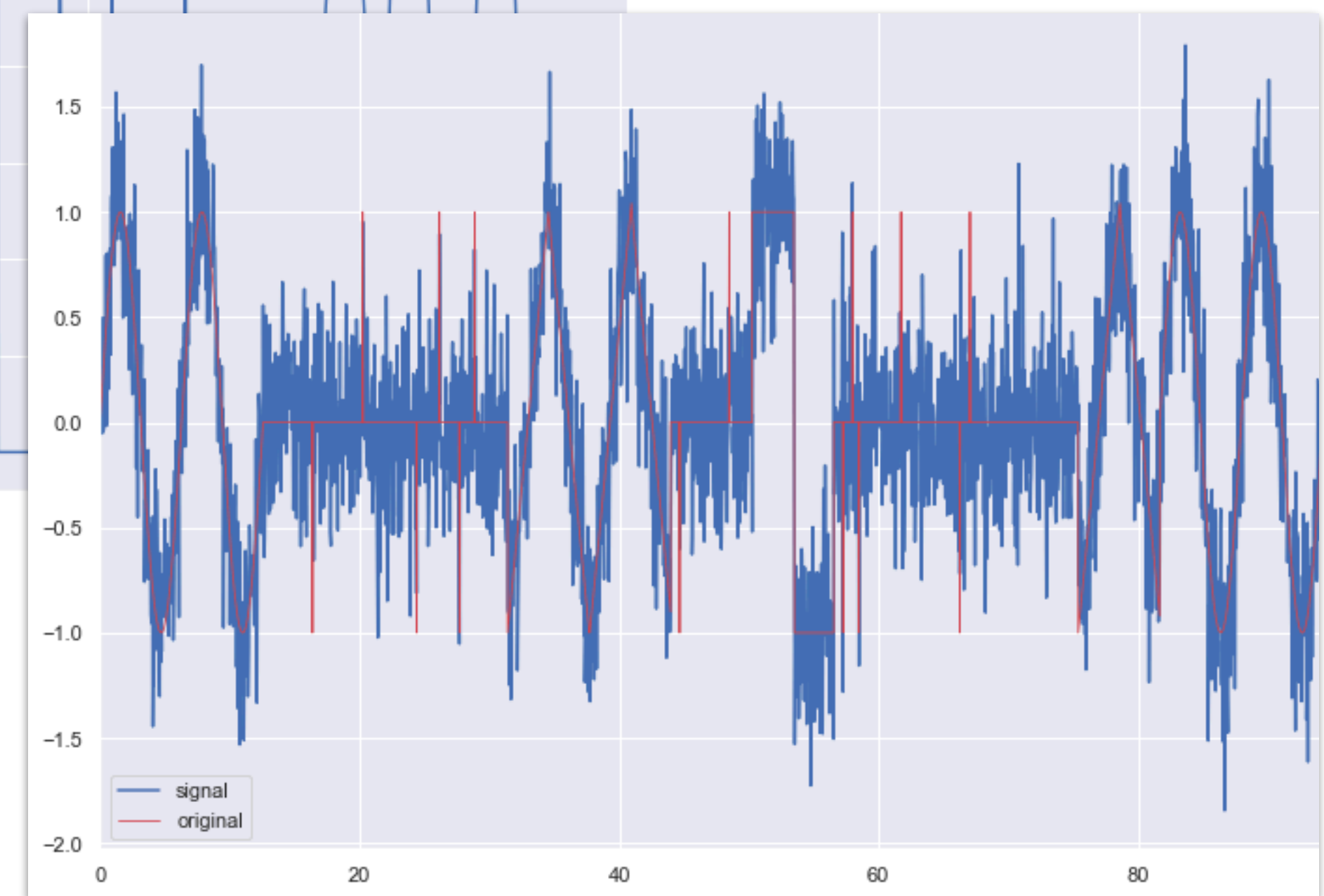
[examples/different-sensors-jsimon/problem-description.ipynb](https://github.com/jsimon123/examples/blob/master/different-sensors-jsimon/problem-description.ipynb)

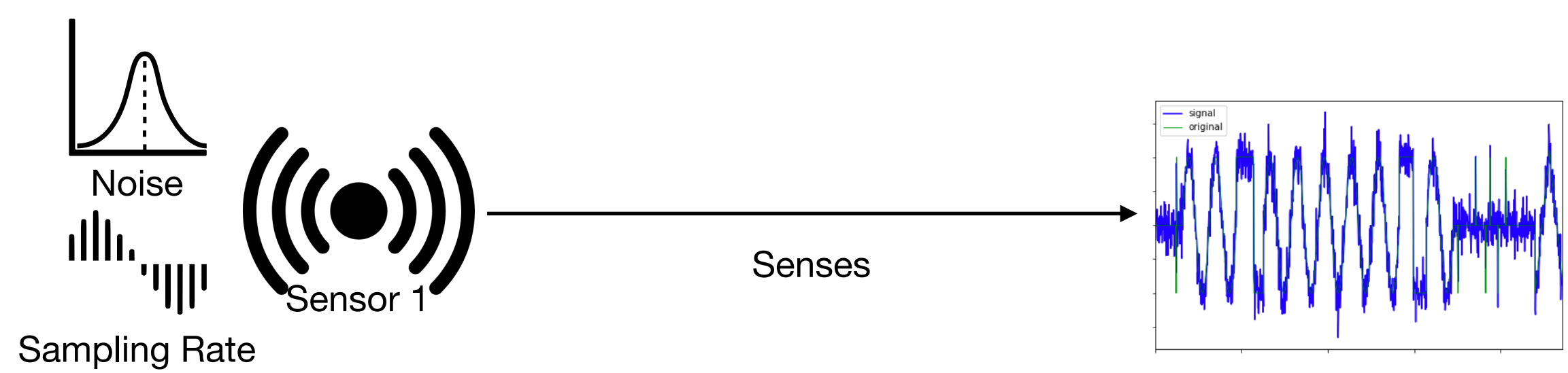
Unifying Sensor-Recordings for later Supervised Training

Sensor 1

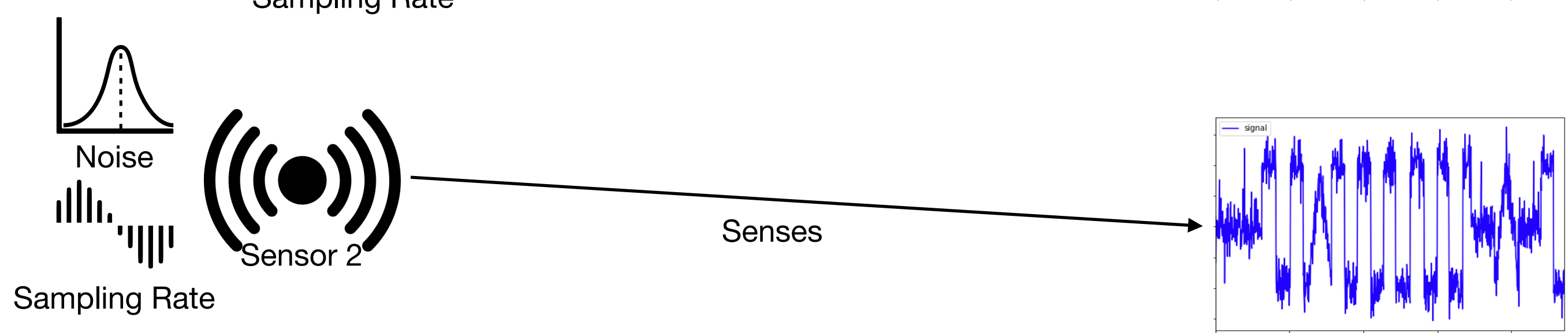


Sensor 2

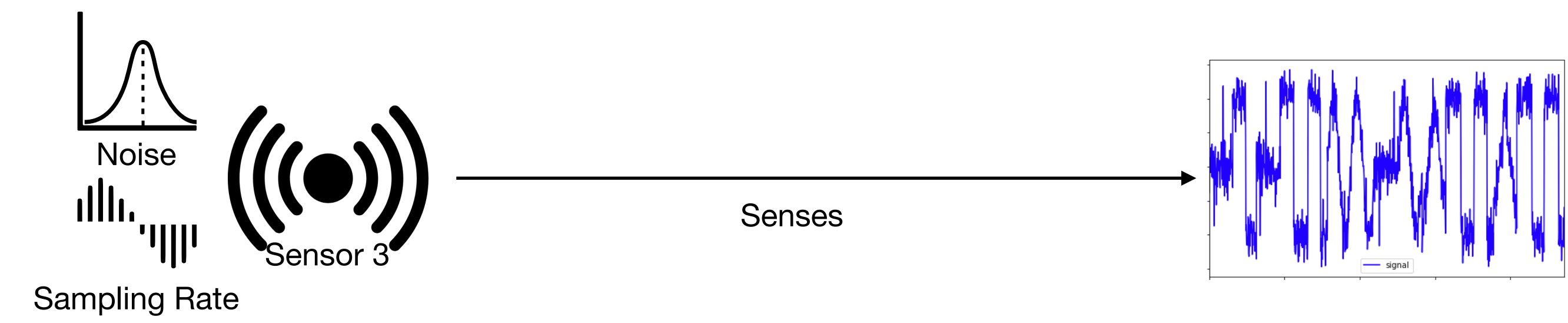




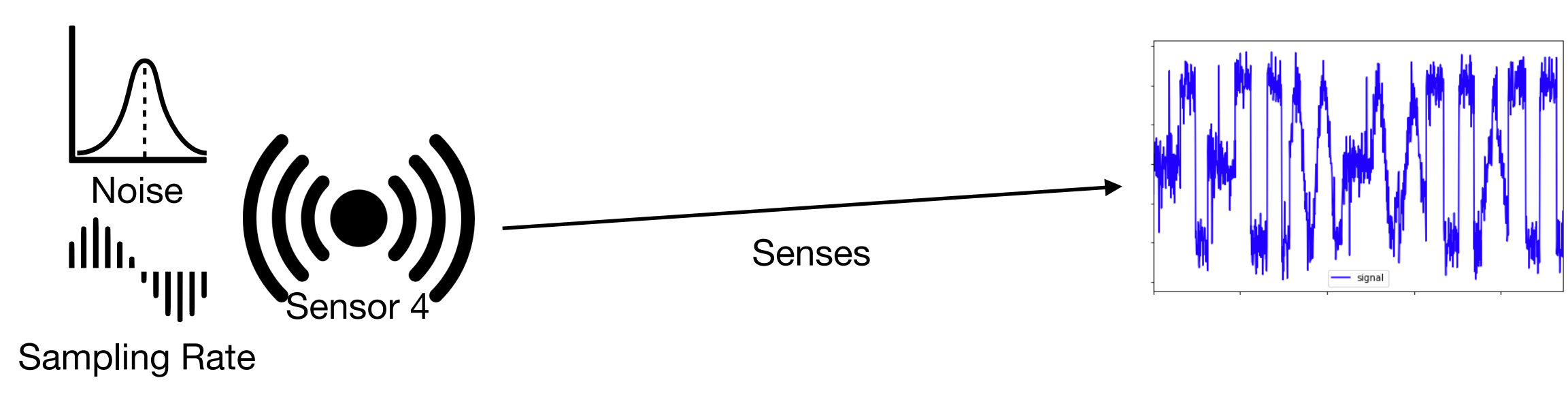
Experiment 1



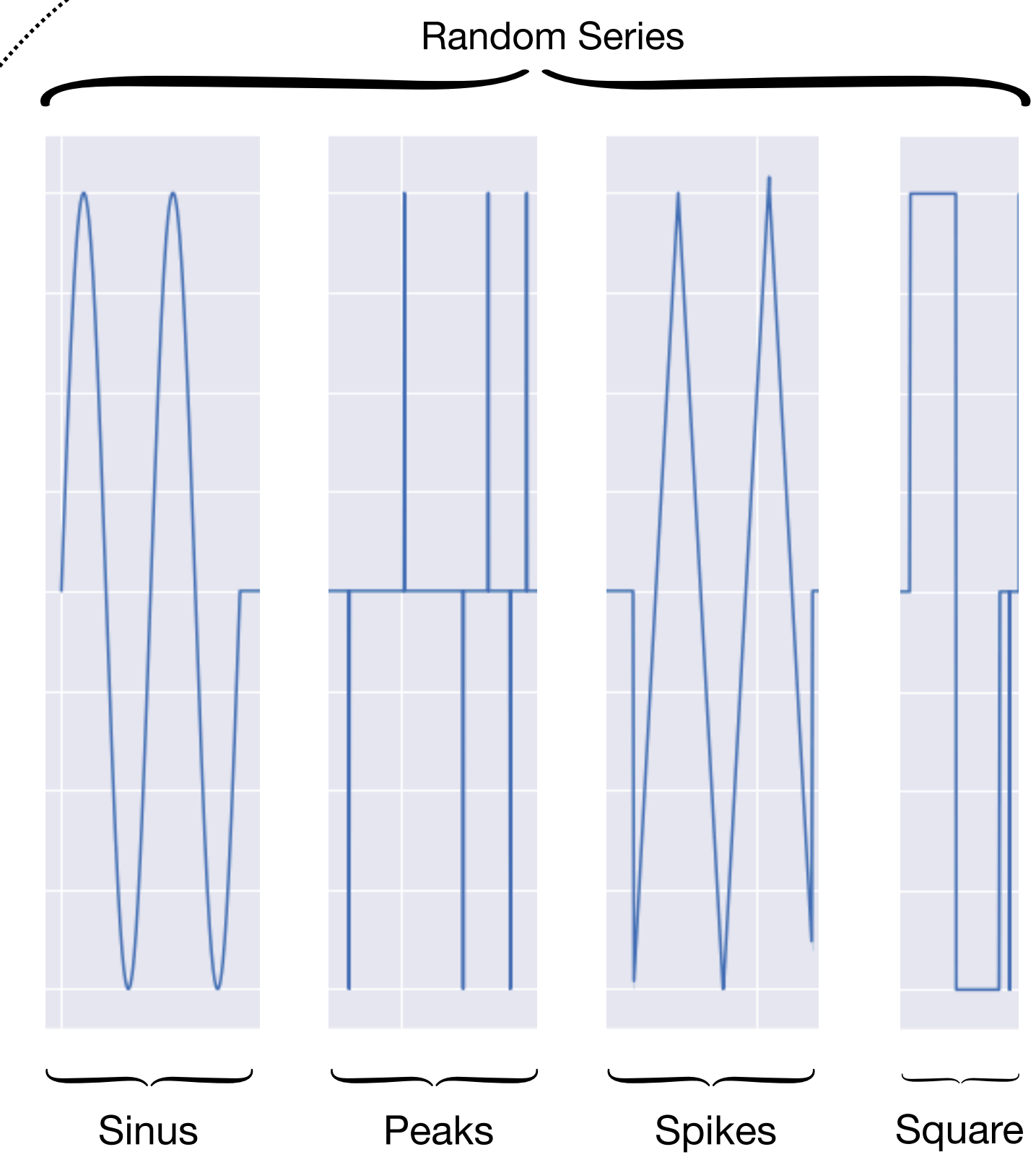
Experiment 2



Experiment 3

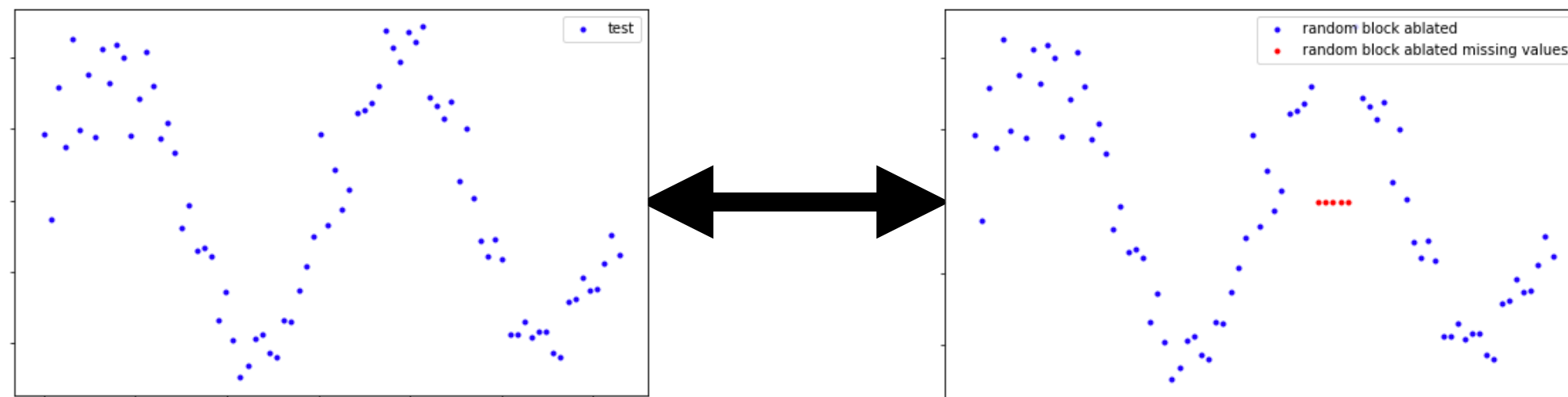


Experiment 4

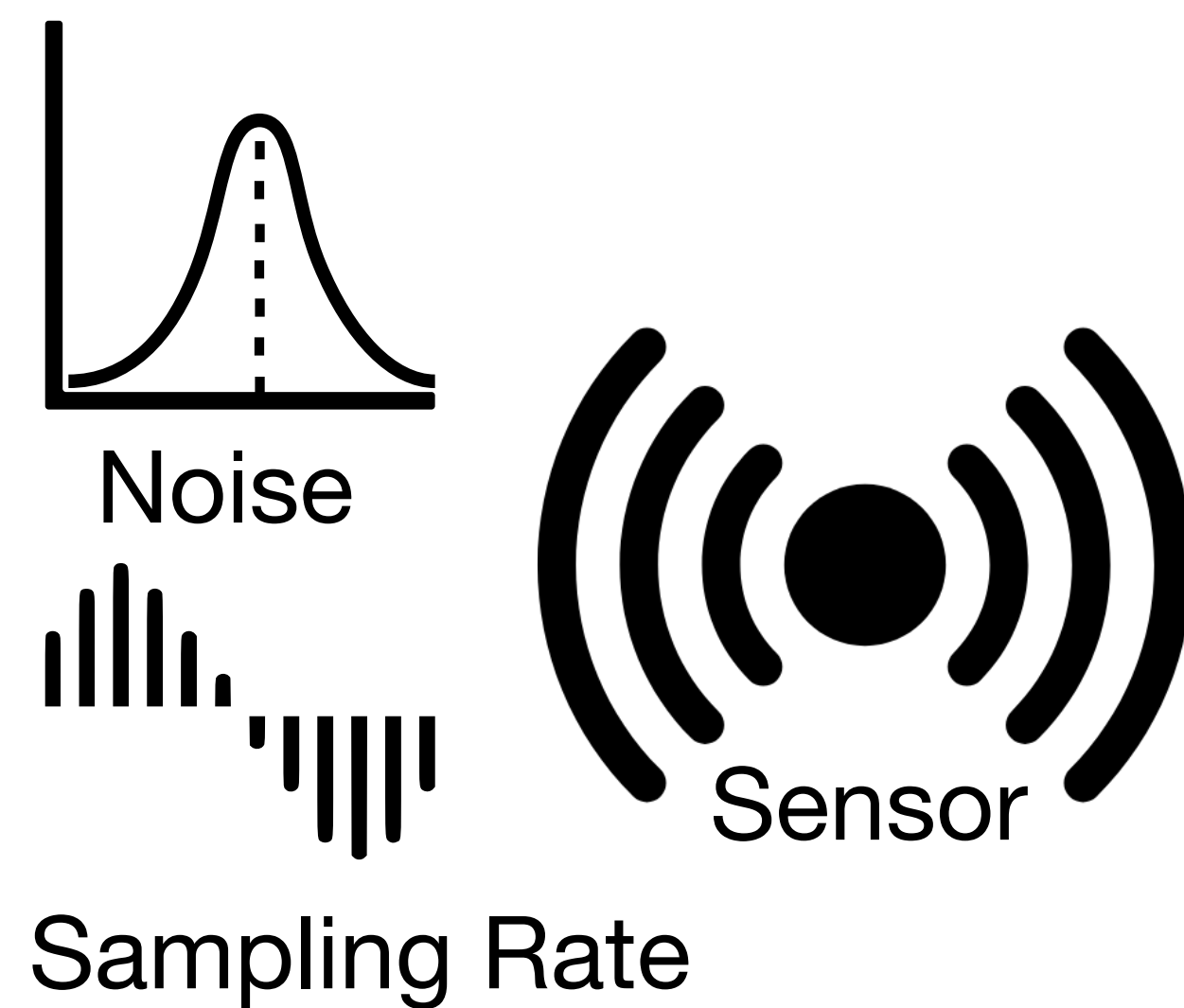


**This means we can
not directly calibrate
sensors**

Because we do not know when they sensed the same



$$\overrightarrow{x} = \begin{bmatrix} 0.932 & 0 \\ 0. & 1 \\ 0.031 & 0 \\ 0. & 1 \\ 0.618 & 0 \end{bmatrix}$$

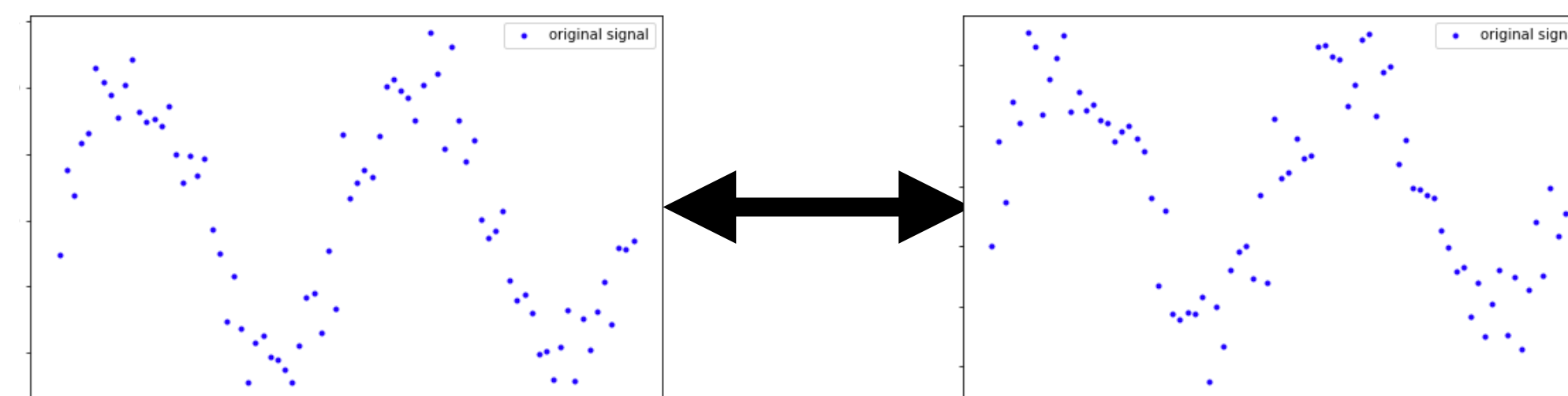
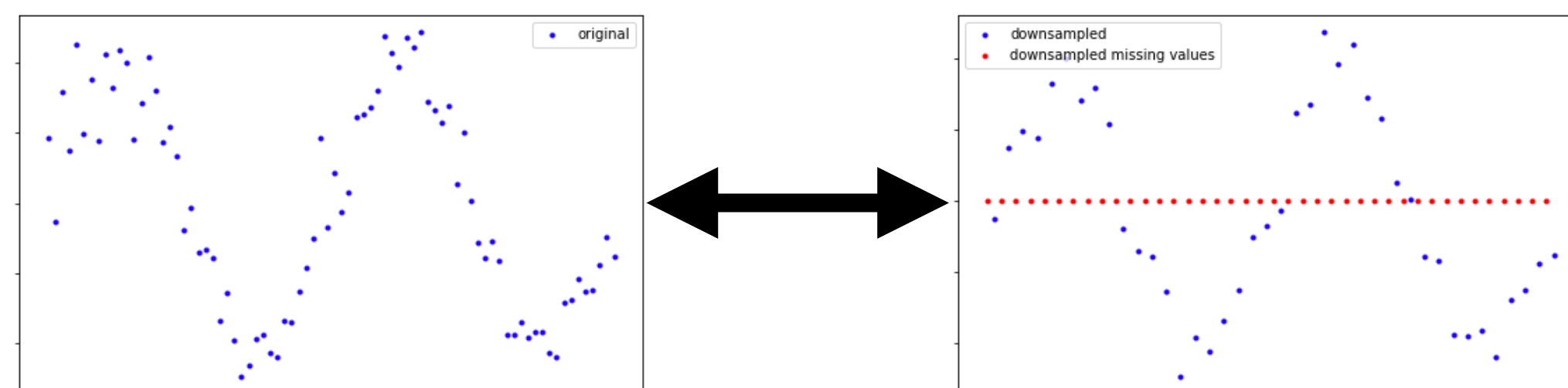


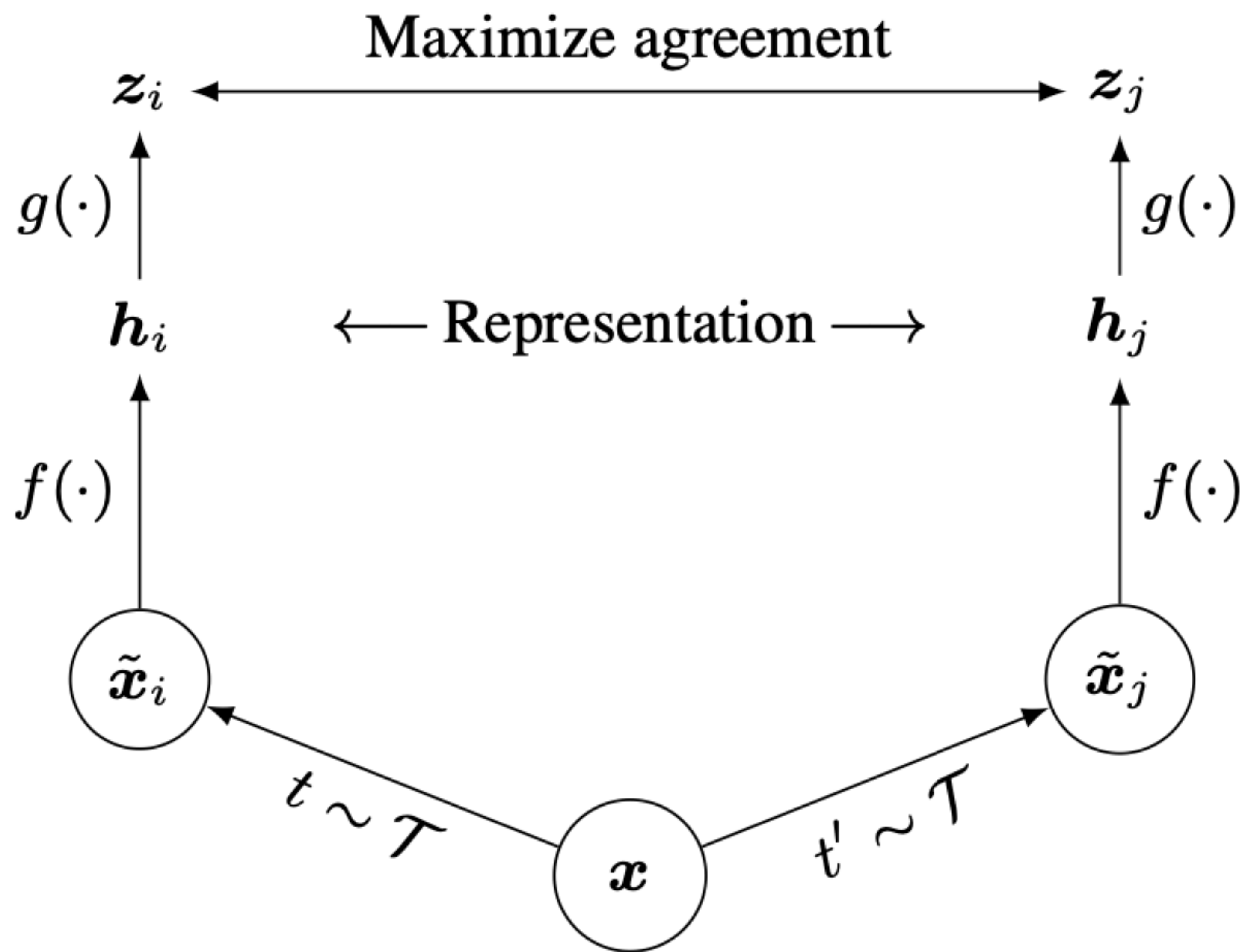
Common Representation (100Hz)

Noise

Sampling Rate

Ablations (sections of sensor failure)





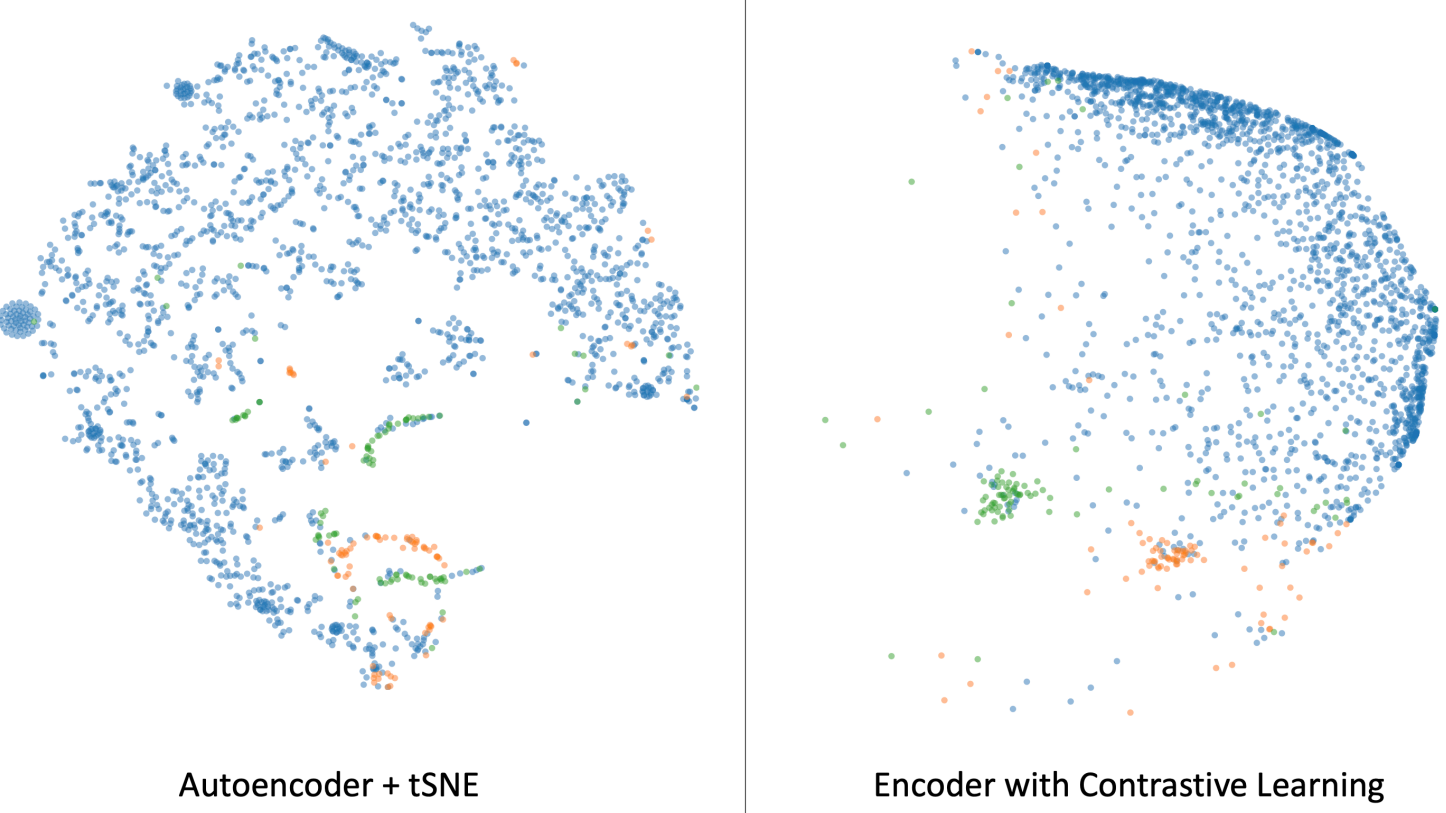
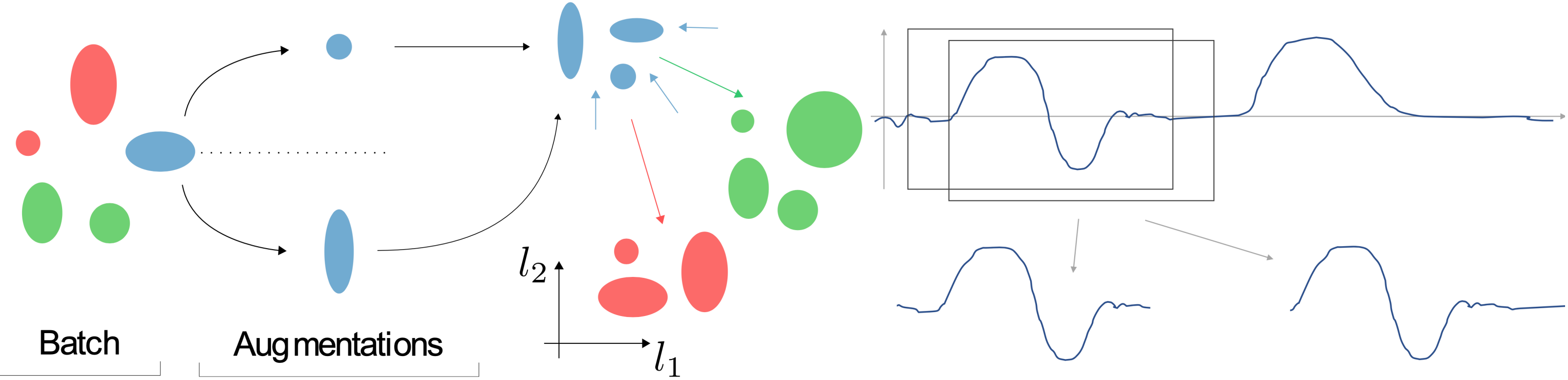
Demo Time

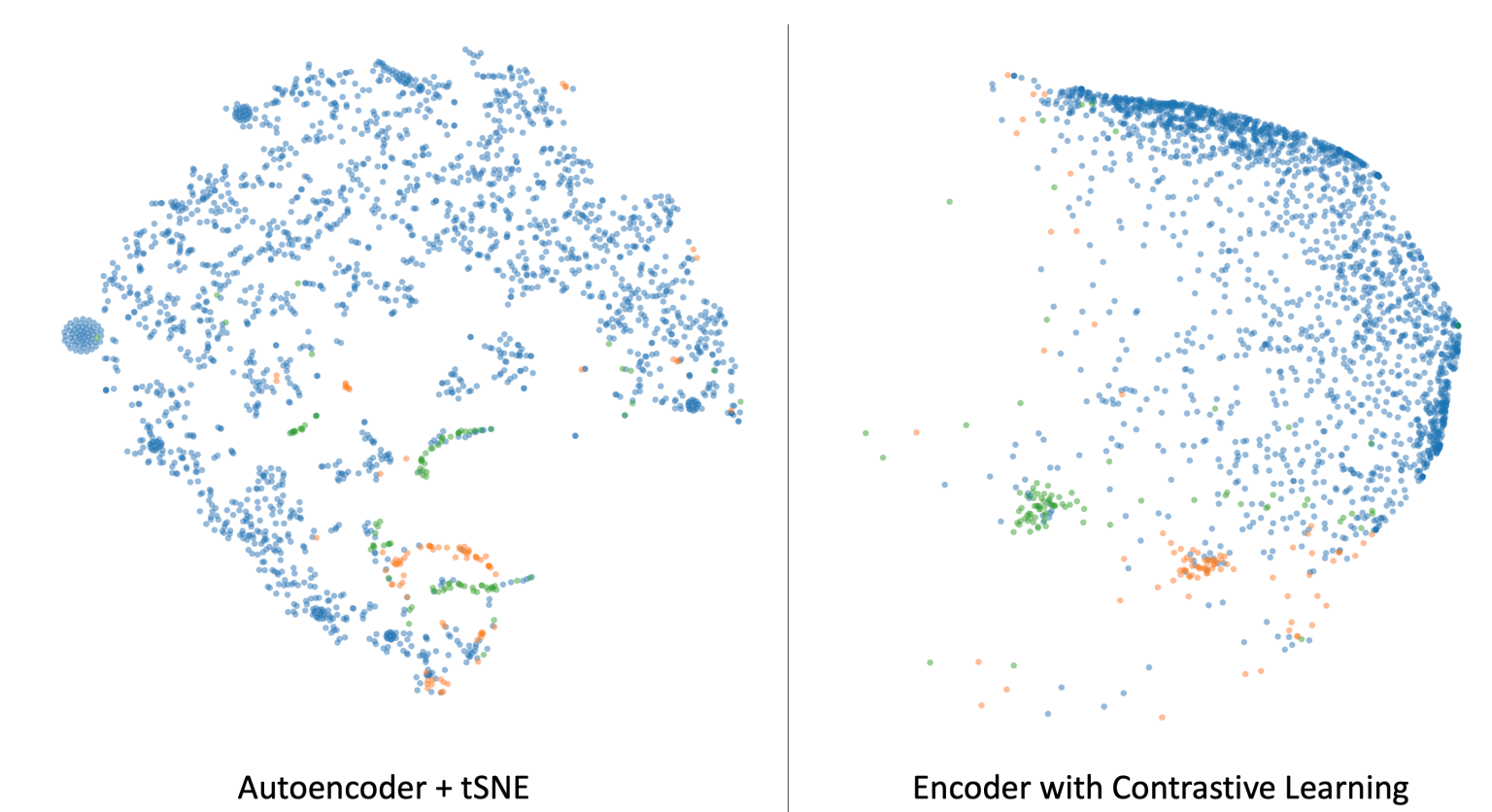
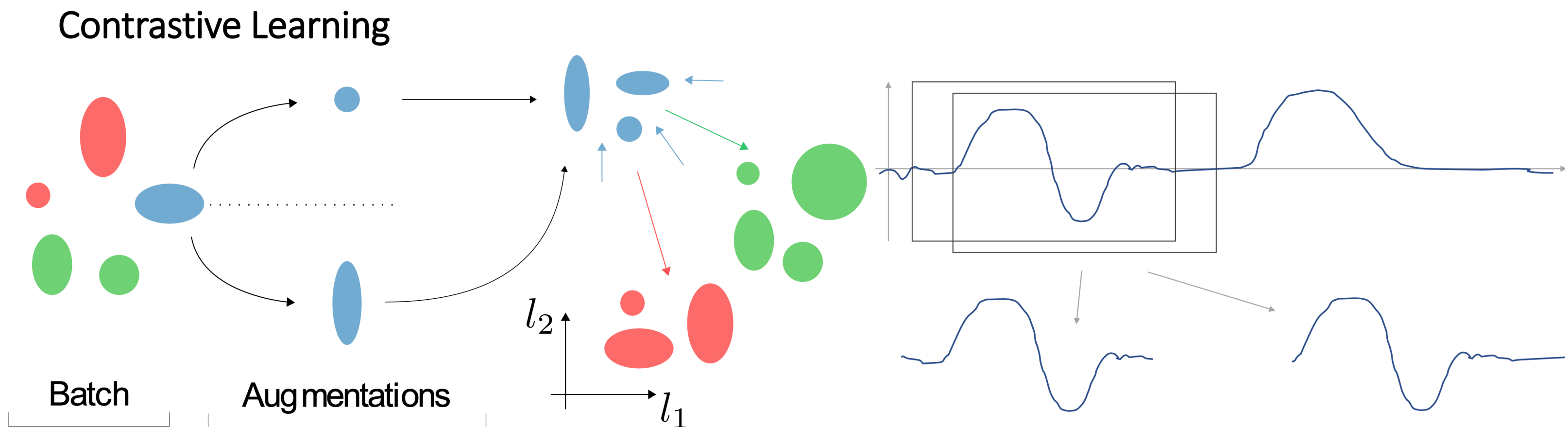
<https://github.com/grazai/contrastive-learning-june-2020>

examples/different-sensors-jsimon

Summary

Contrastive Learning





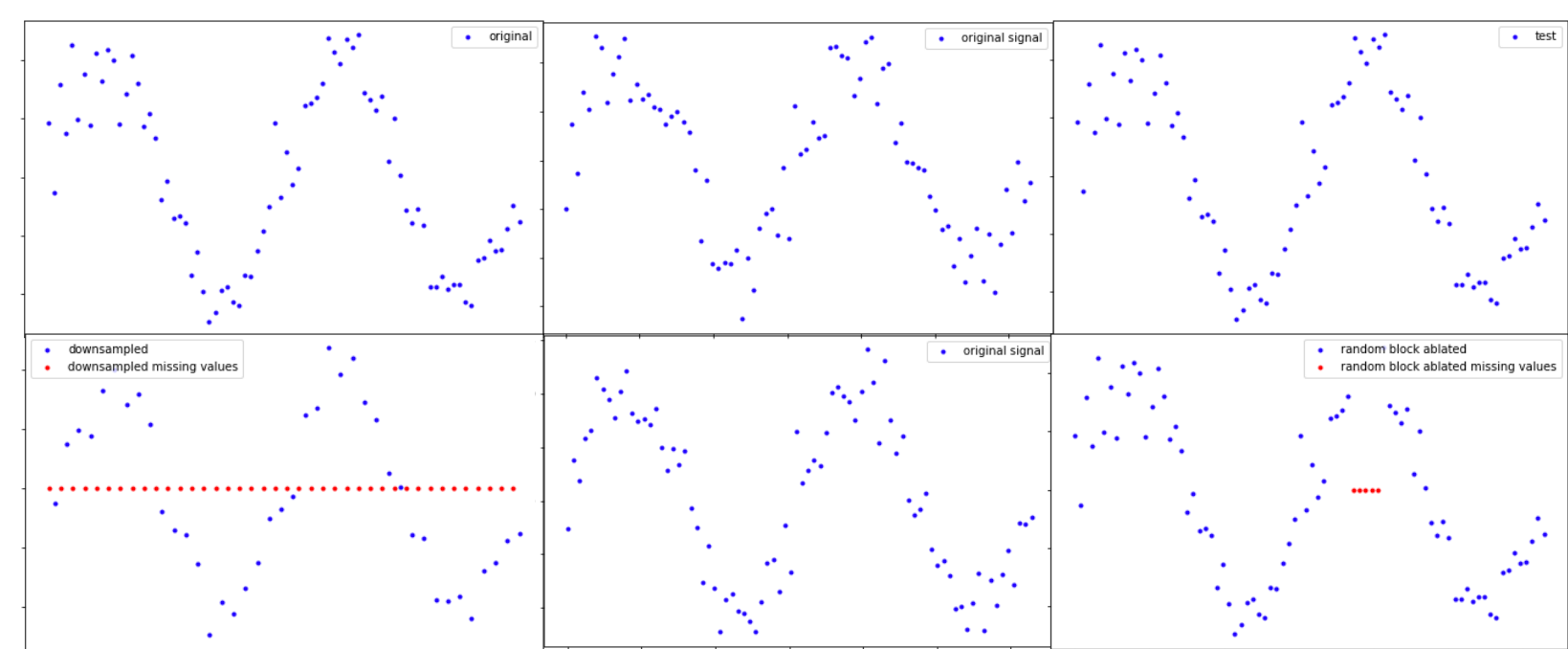
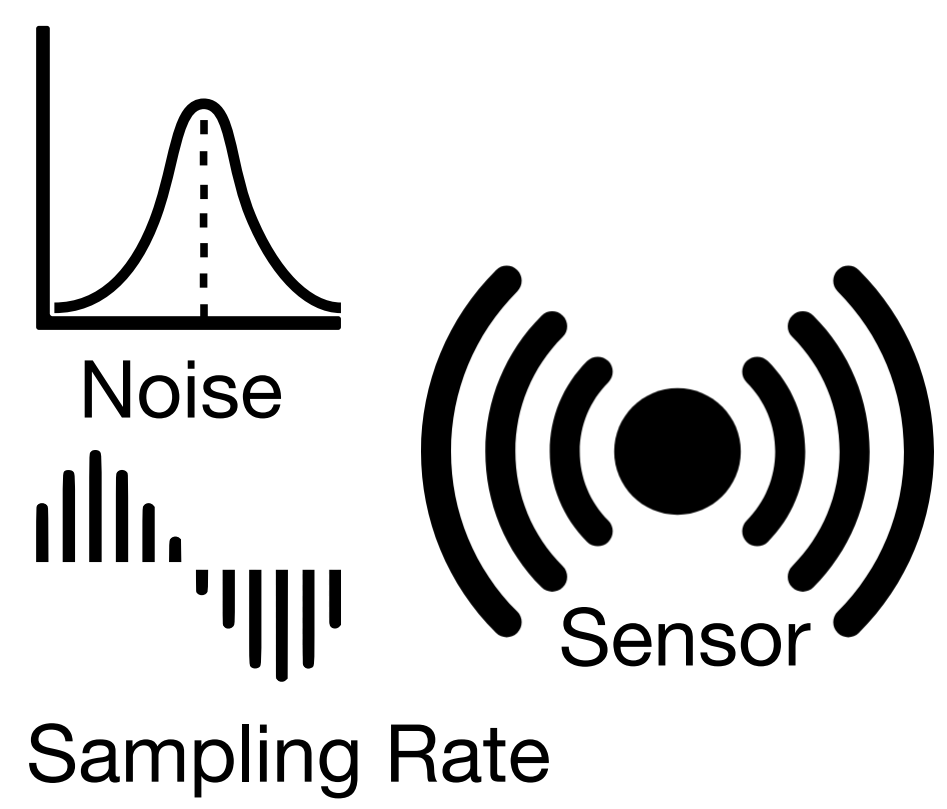
2005



2012



2020



Thanks for the Attention

We will continue for an Q&A
in discord (we will post the
link in the twitch chat)

