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

- little 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

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

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

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

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

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

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

Contrastive Loss

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

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

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

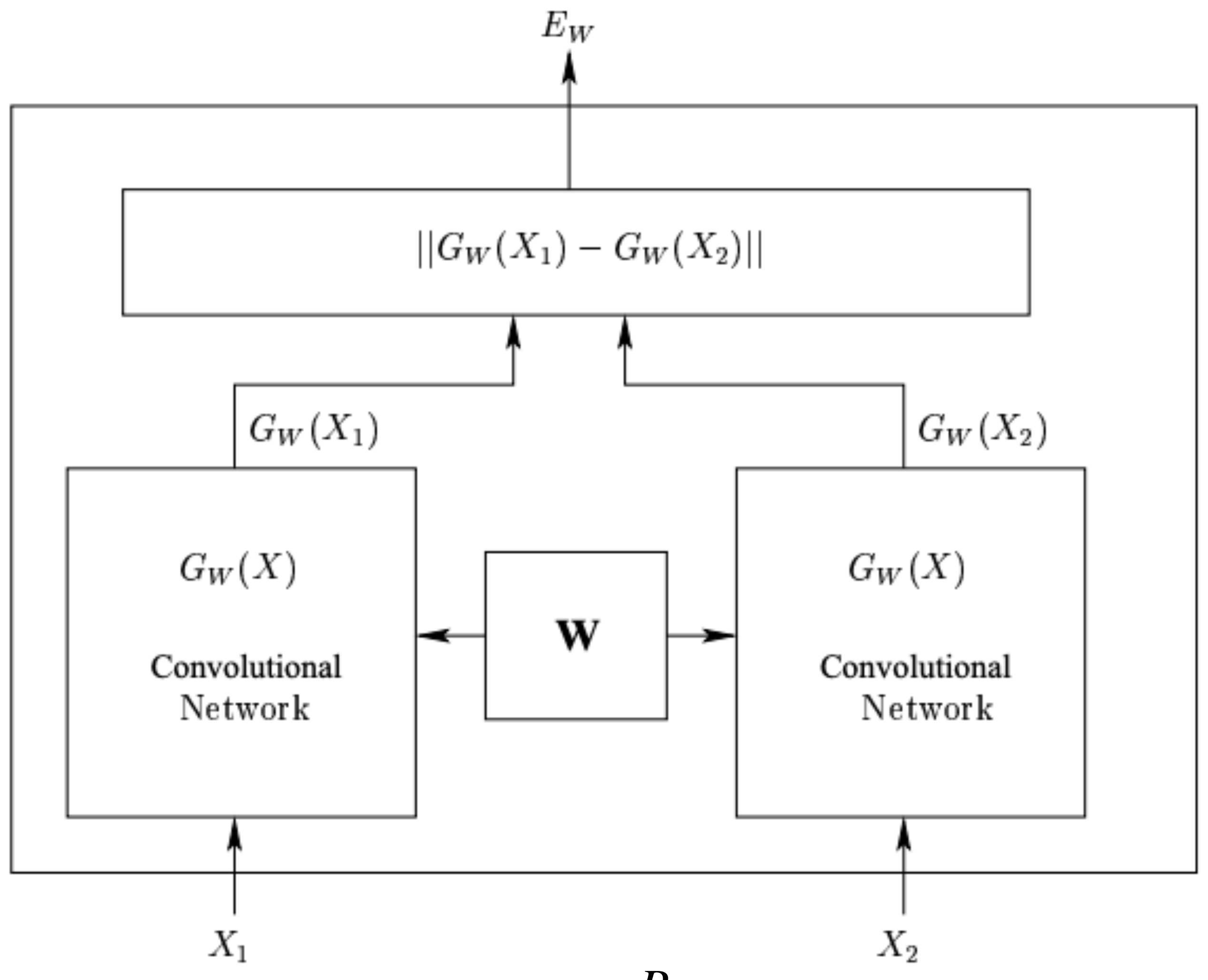
imposter pair

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

Siamese Networks



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

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


Sumit Chopra


Raia Hadsell


Yann LeCun

2005 CVPR: Face Verification

imposter pair genuine pair

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

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

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

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*

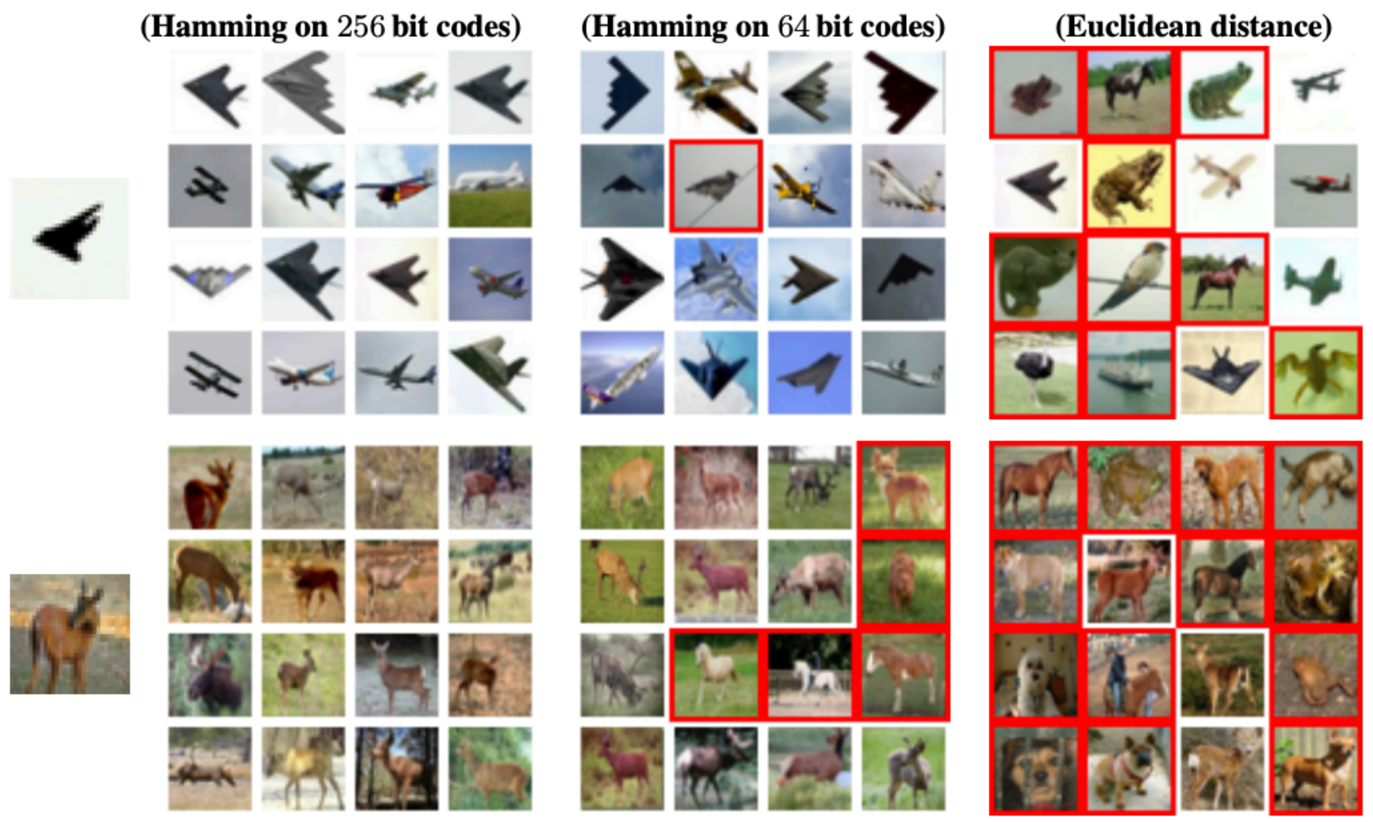
History and other Terms for the same


Mohammad Norouzi


David J. Fleet


Ruslan Salakhutdinov

2012 NeurIPS: Distances on Images



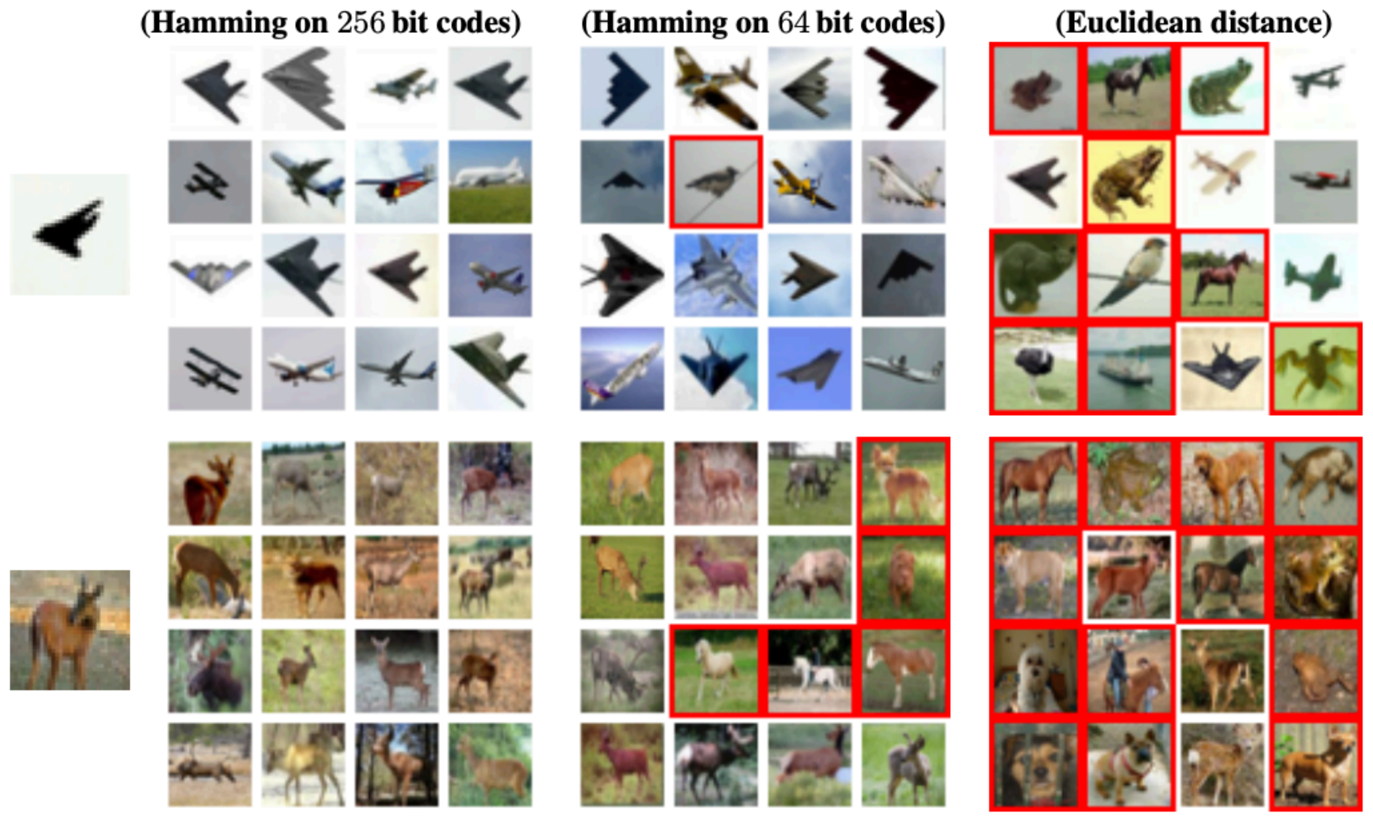
Triplet Loss

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

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

Metric Learning



2012 NeurIPS: Distances on Images

Mohammad Norouzi

David J. Fleet

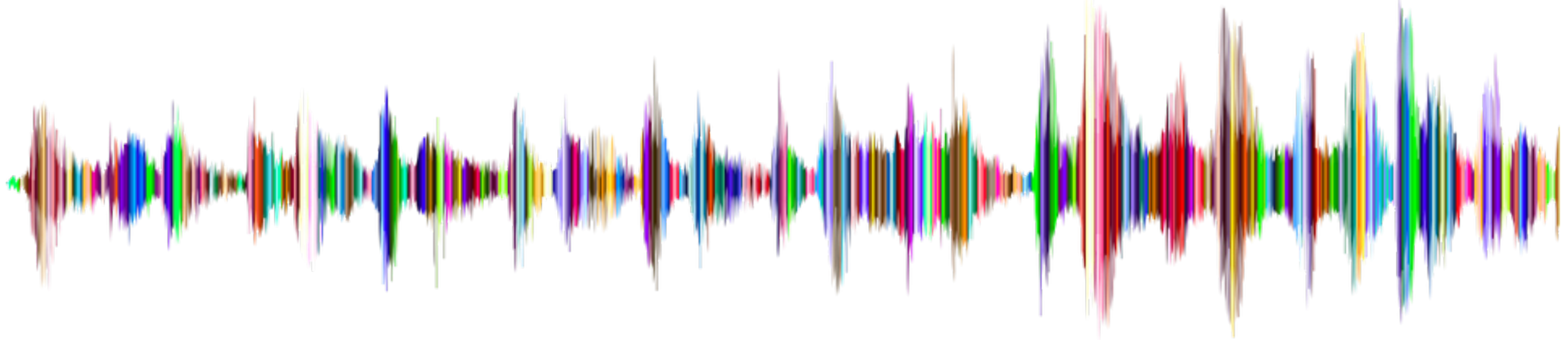
Ruslan Salakhutdinov

Triplet Loss

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

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



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

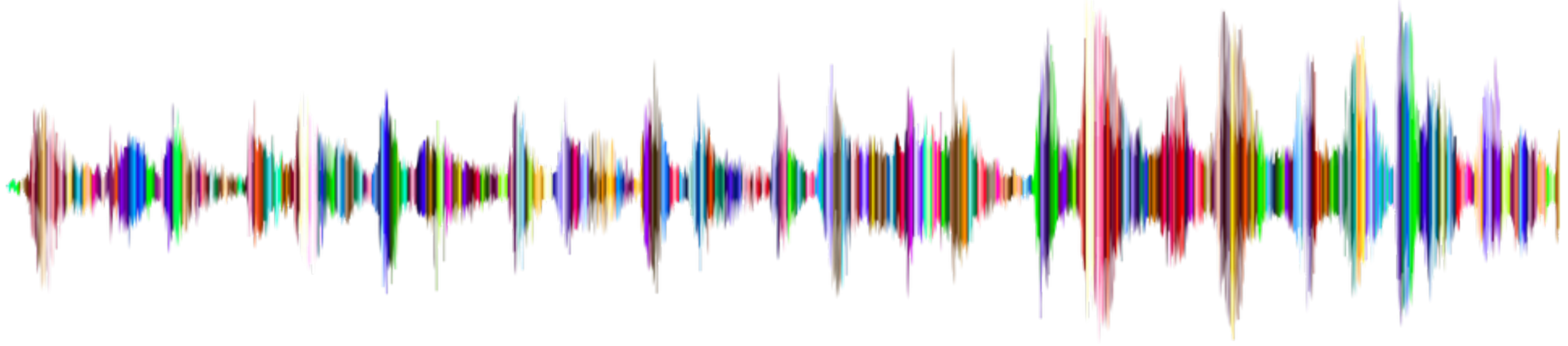
2018 Application to Sensor Data

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**

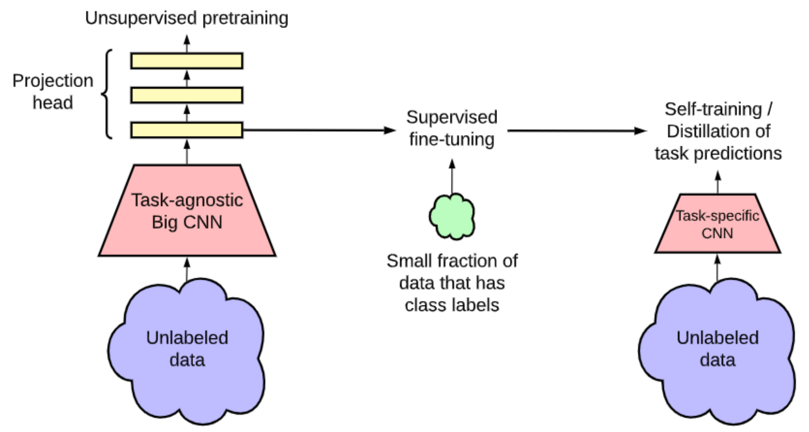
Srinivas, A., Laskin, M., Abbeel, P. (2020). **CURL: Contrastive Unsupervised Representations for Reinforcement Learning** arXiv https://arxiv.org/abs/2004.04136

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

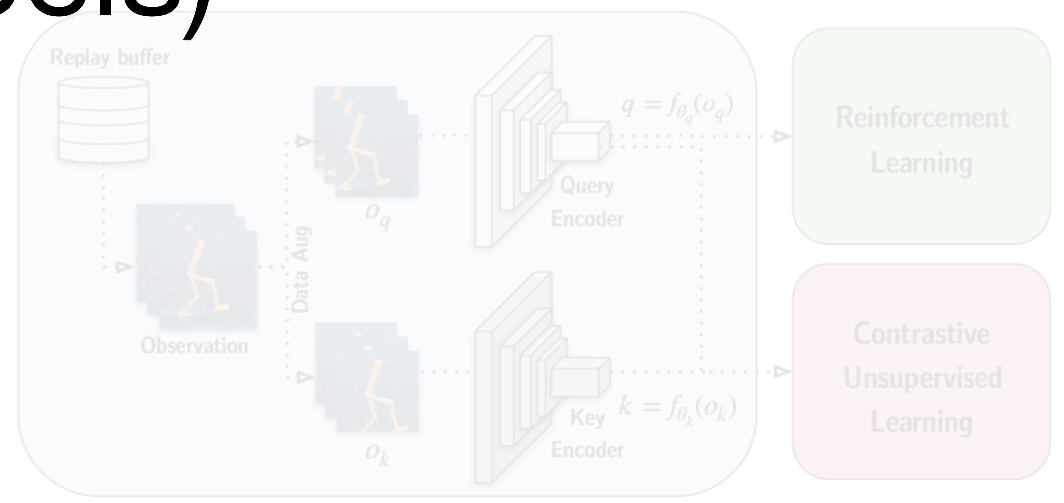
Chen, T., Kornblith, S., Norouzi, M., Hinton, G. (2020). **A Simple Framework for Contrastive Learning of Visual Representations**https://arxiv.org/abs/2002.05709



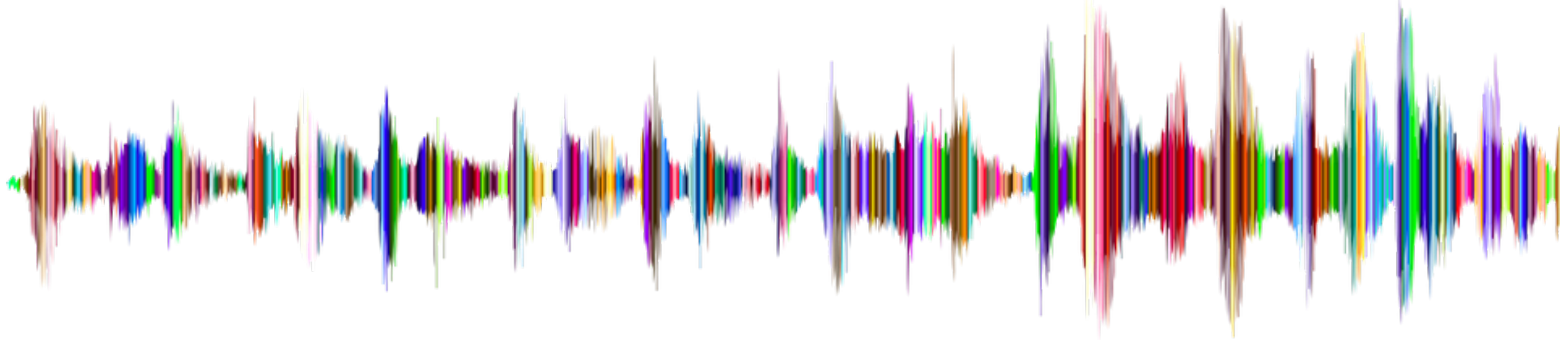
Chen, T., Kornblith, S., Swersky, K., Norouzi, M., Hinton, G. (2020). **Big Self-Supervised Models are Strong Semi-Supervised Learners**

Early 2020 Sample Efficiency (little labels)

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

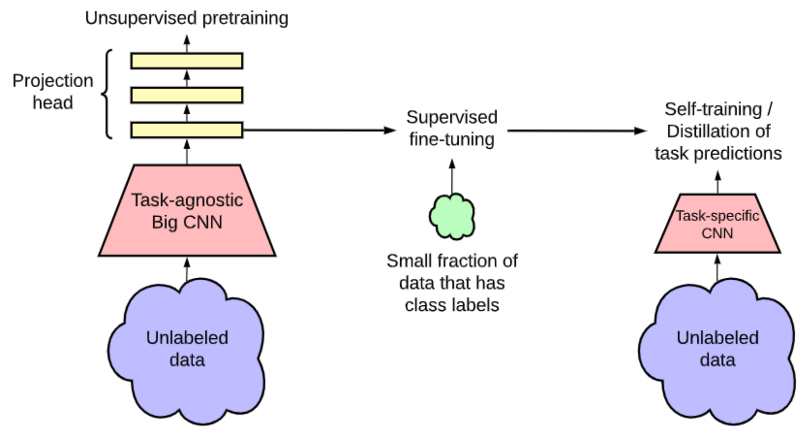


History and other Terms for the same



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

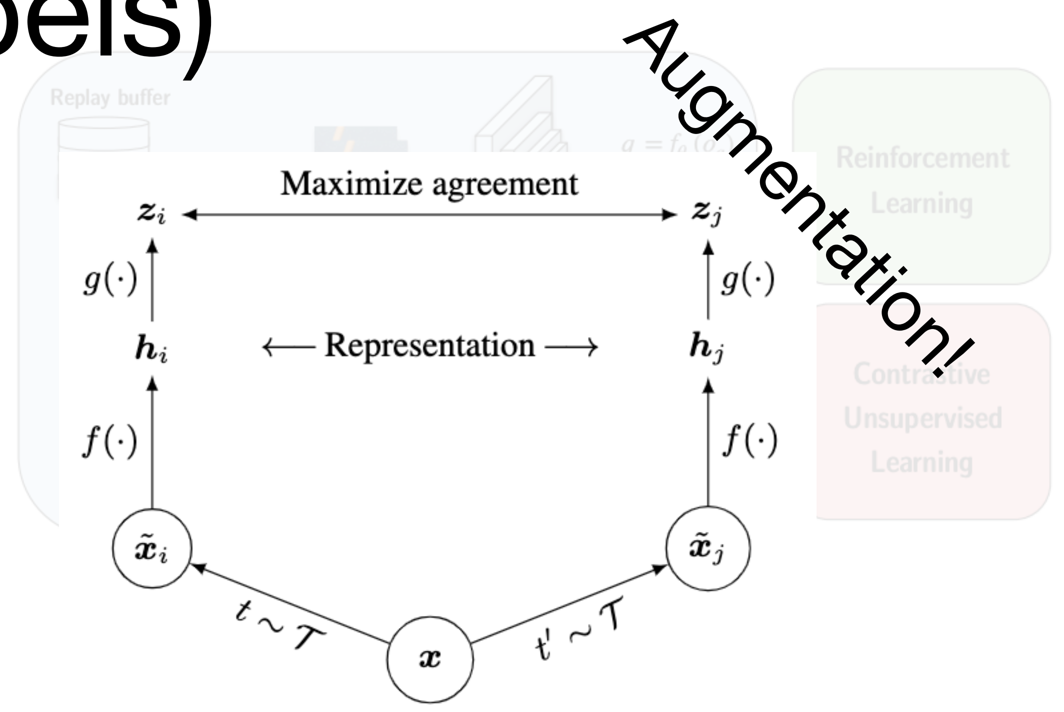


Chen, T., Kornblith, S., Swersky, K., Norouzi, M., Hinton, G. (2020). **Big Self-Supervised Models are Strong Semi-Supervised Learners**

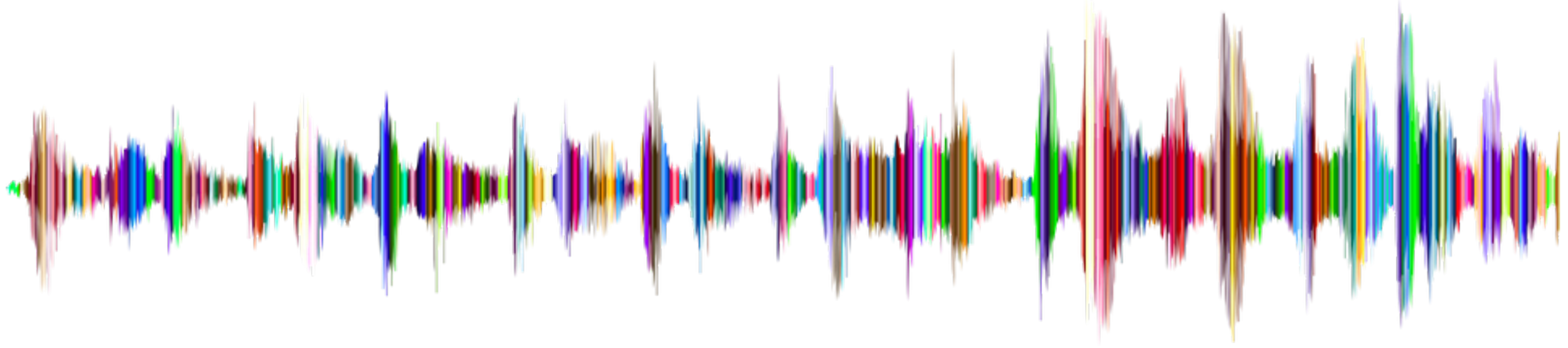
Early 2020 Sample Efficiency (little labels)

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

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

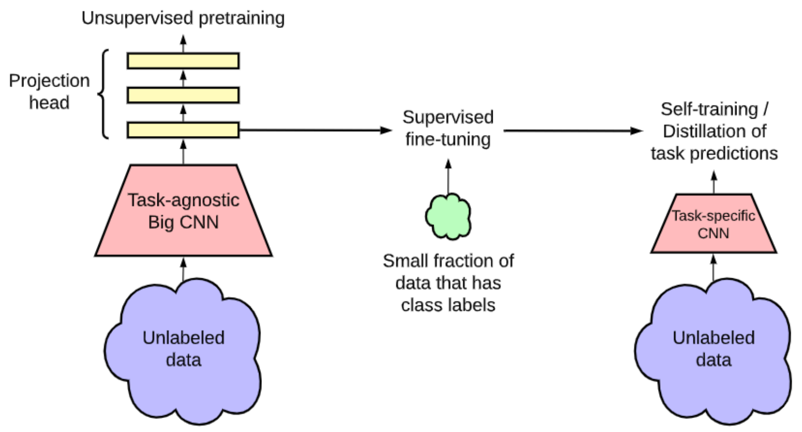


History and other Terms for the same



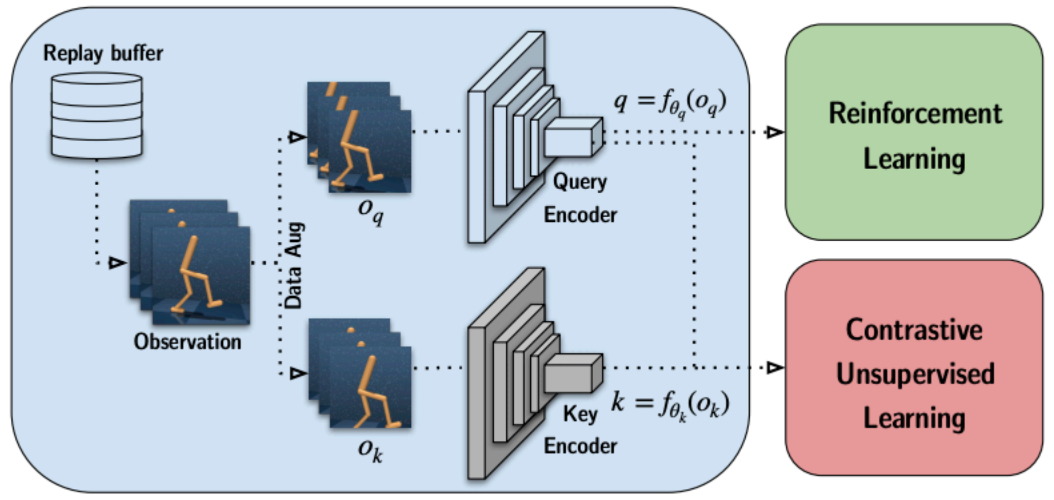
Lian, Z., Li, Y., Tao, J., Huang, J. (2018). **Speech Emotion Recognition via Contrastive Loss under Siamese Networks**https://dx.doi.org/10.1145/3267935.3267946

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**

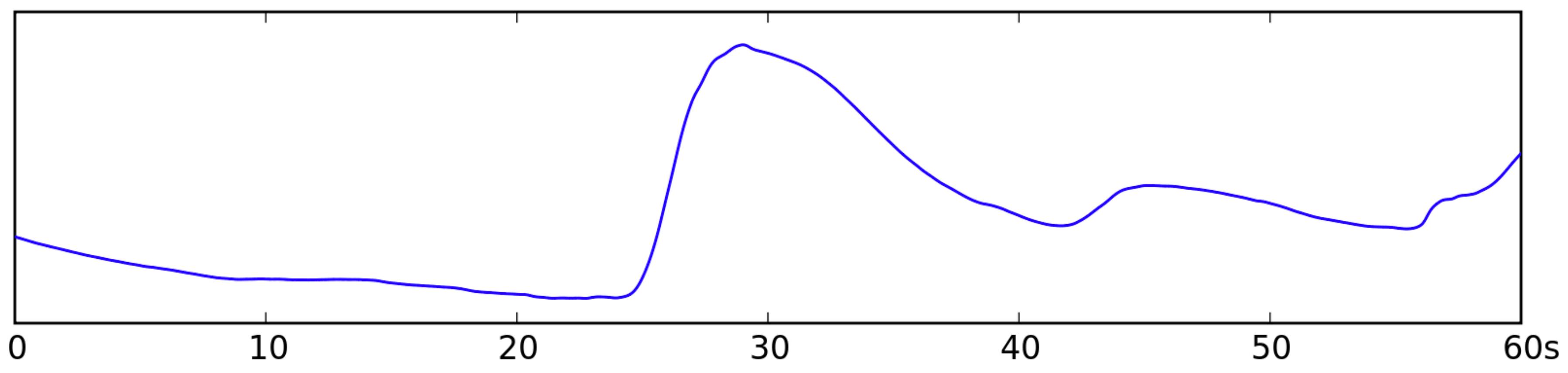
Srinivas, A., Laskin, M., Abbeel, P. (2020). **CURL: Contrastive Unsupervised Representations for Reinforcement Learning** arXiv https://arxiv.org/abs/2004.04136



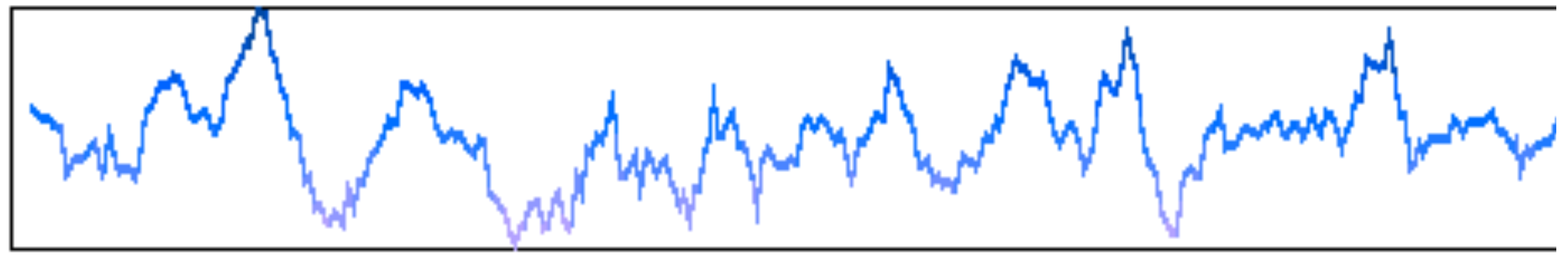
2020 Sample Efficiency in Reinforcement Learning

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

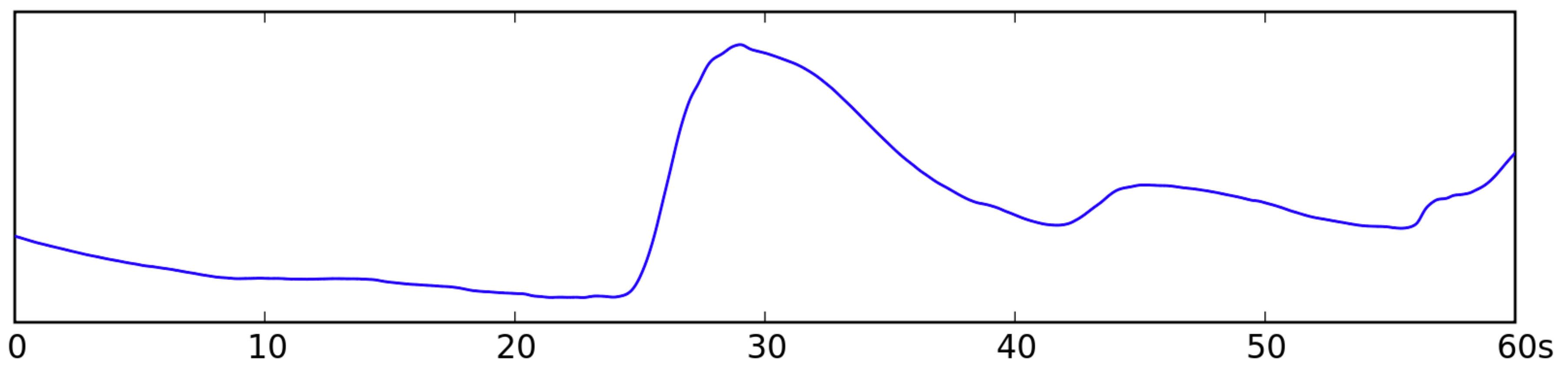


GSR Signal Sensor 2

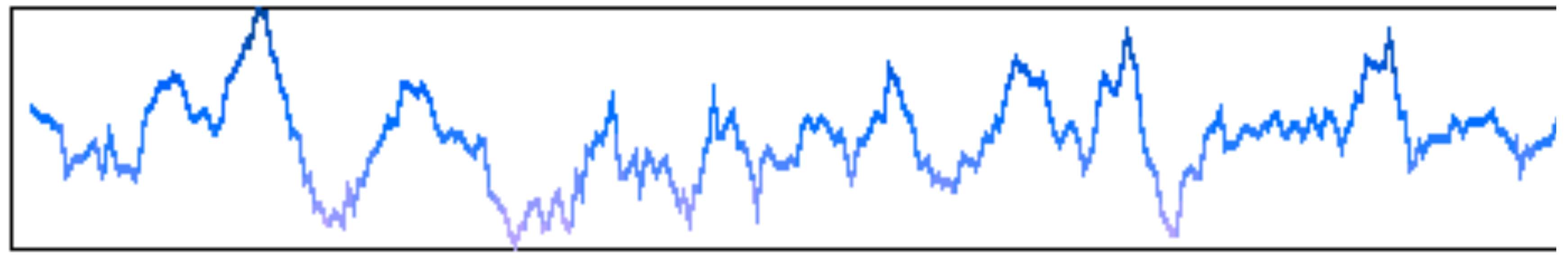
Much more noise!

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



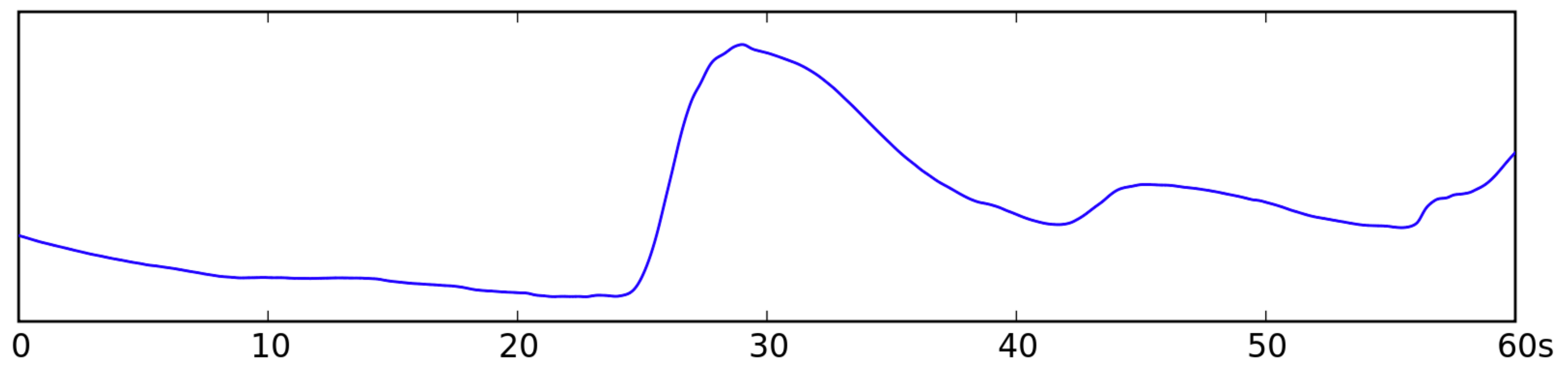
GSR Signal Sensor 2

Much more noise!

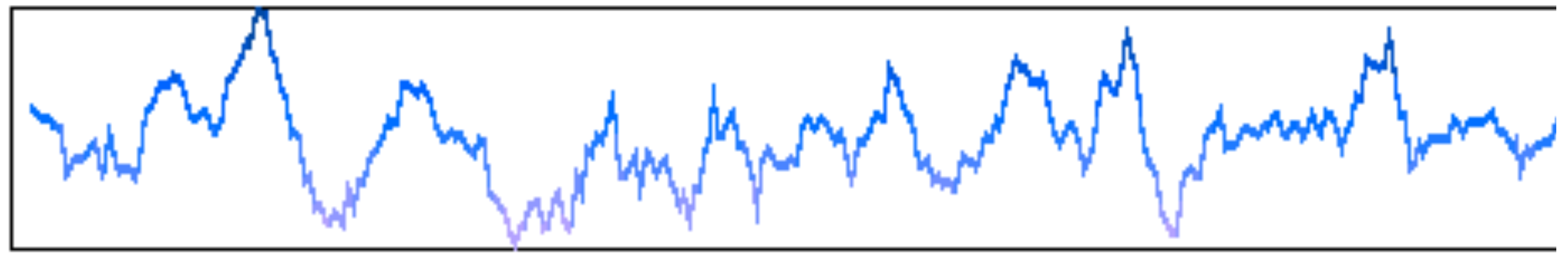
Fractioned Datasets of Drowsiness

- Bad news: I can not show you exactly this use case (confidential datasets)

- Good news: I have a “surrogate problem”



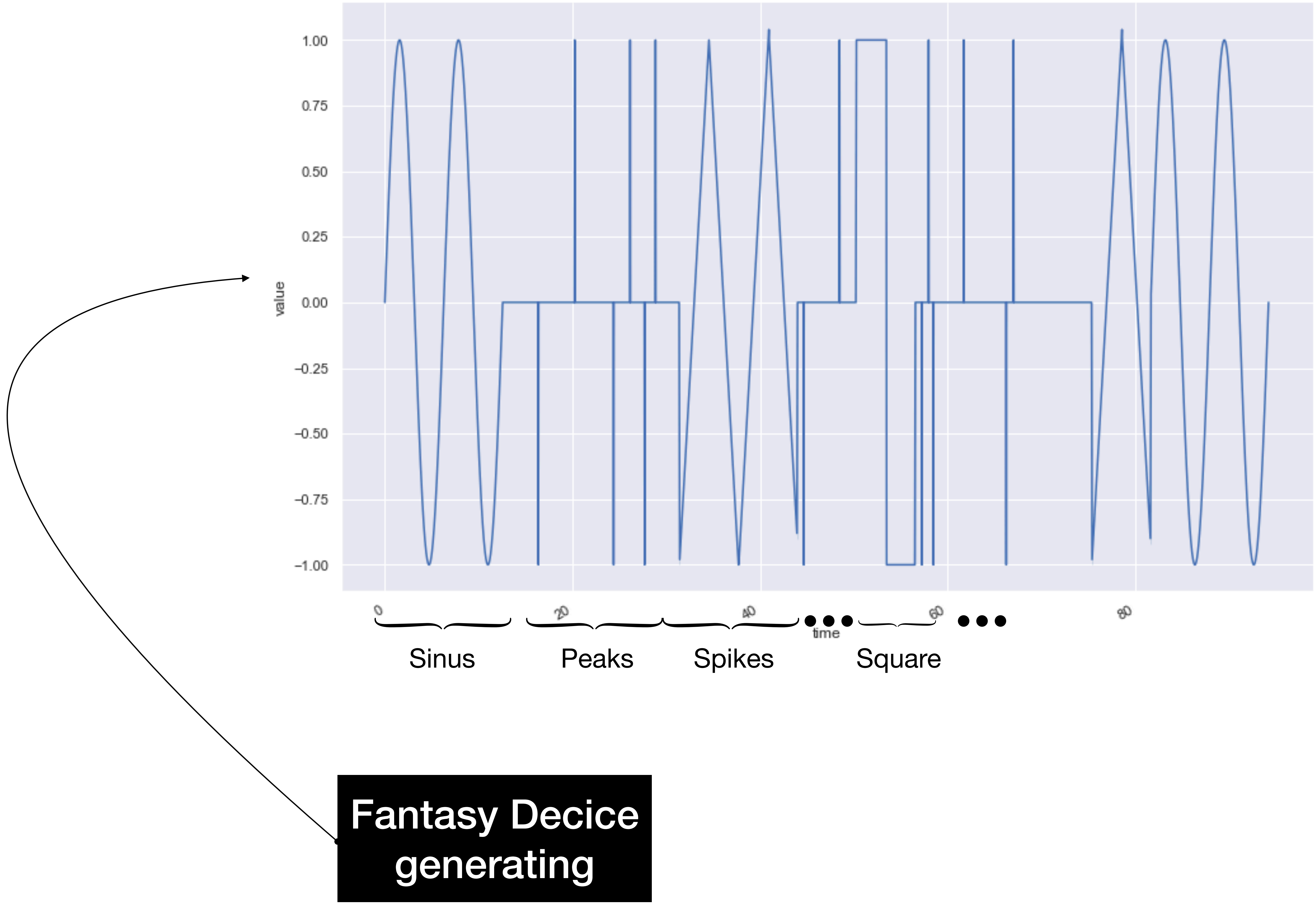
GSR Signal Sensor 1



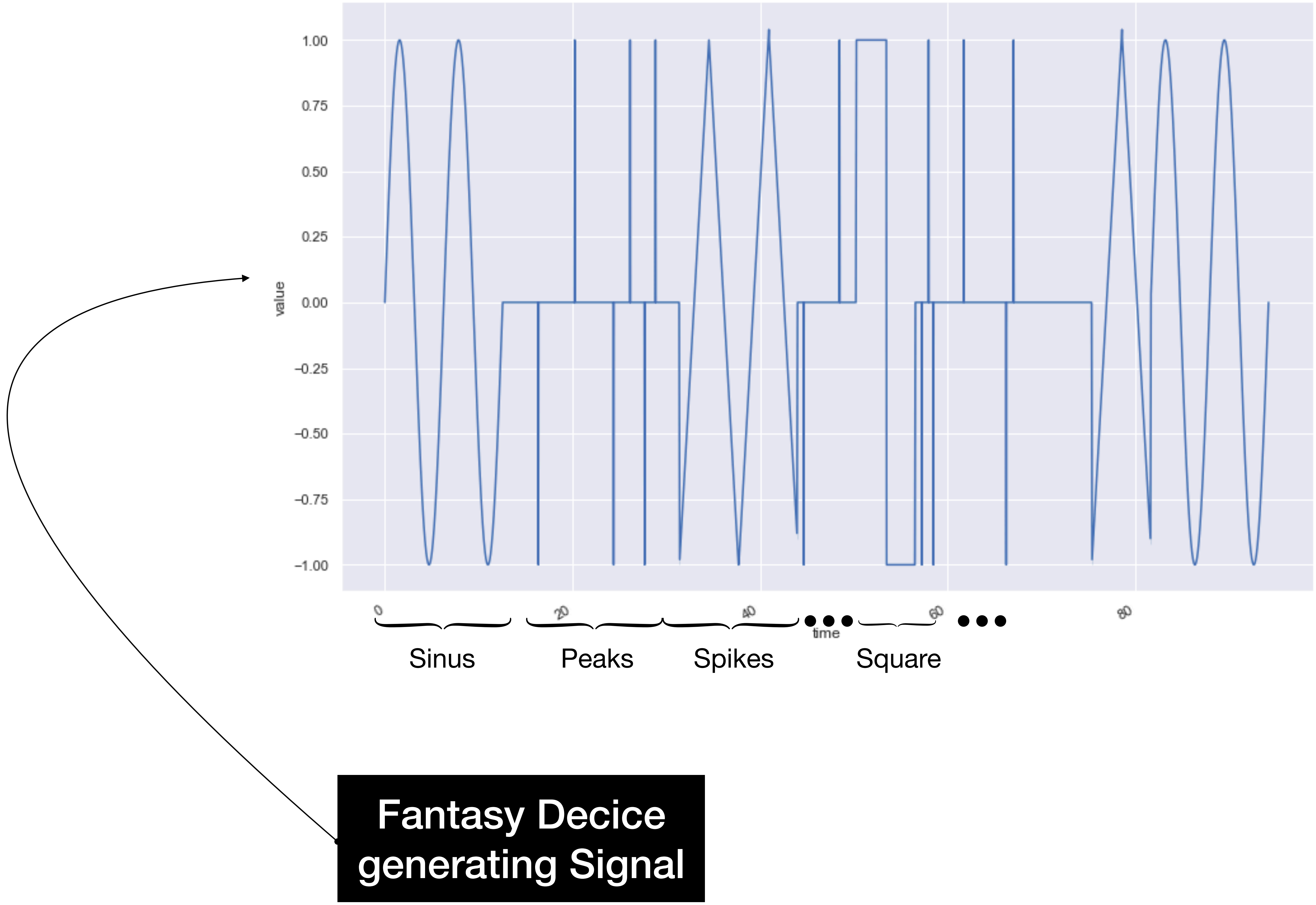
GSR Signal Sensor 2

Much more noise!

Unifying Sensor-Recordings for later Supervised Training



Unifying Sensor-Recordings for later Supervised Training

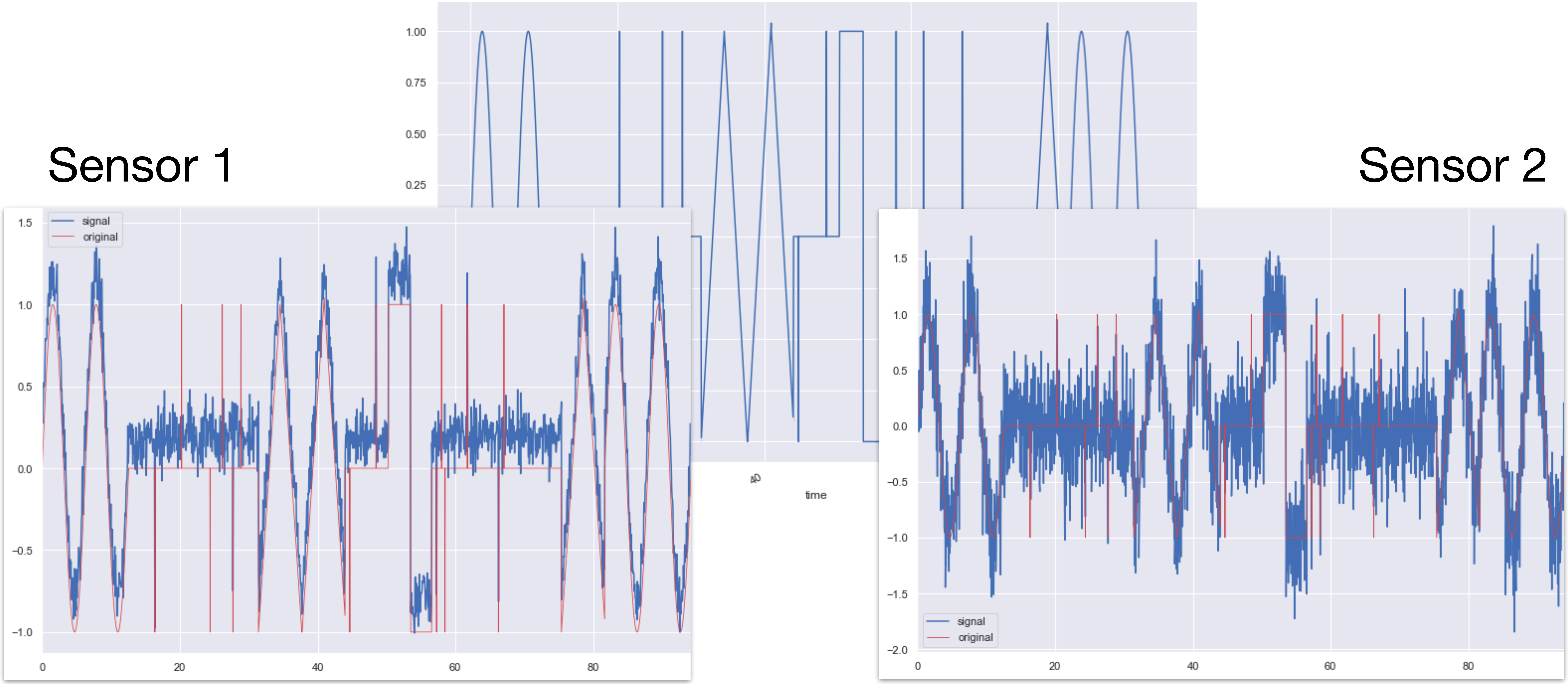


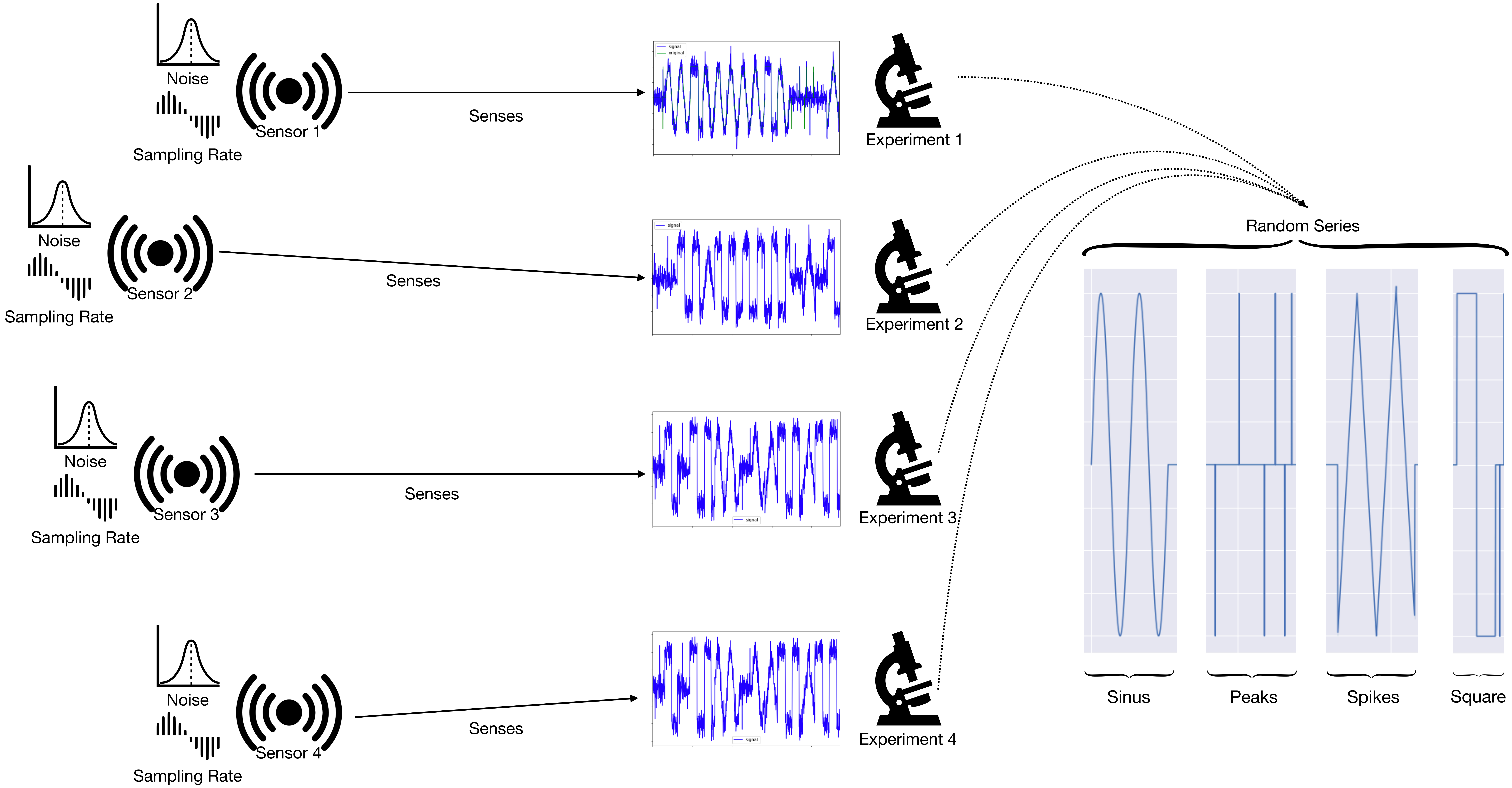
examples/different-sensors-jsimon/problem-description.ipynb

Unifying Sensor-Recordings for later Supervised Training

Sensor 1

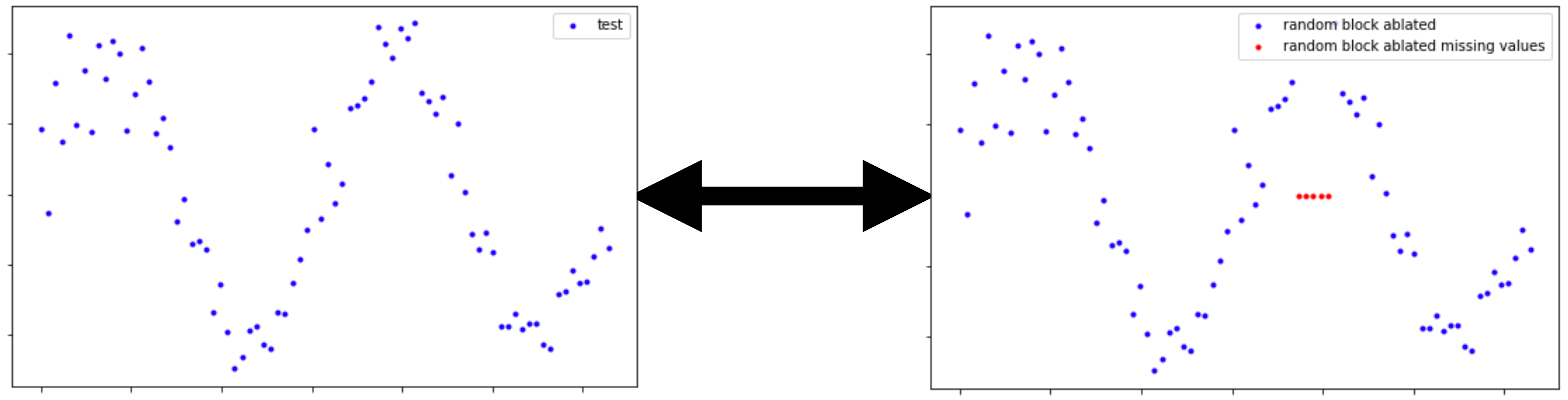
Sensor 2



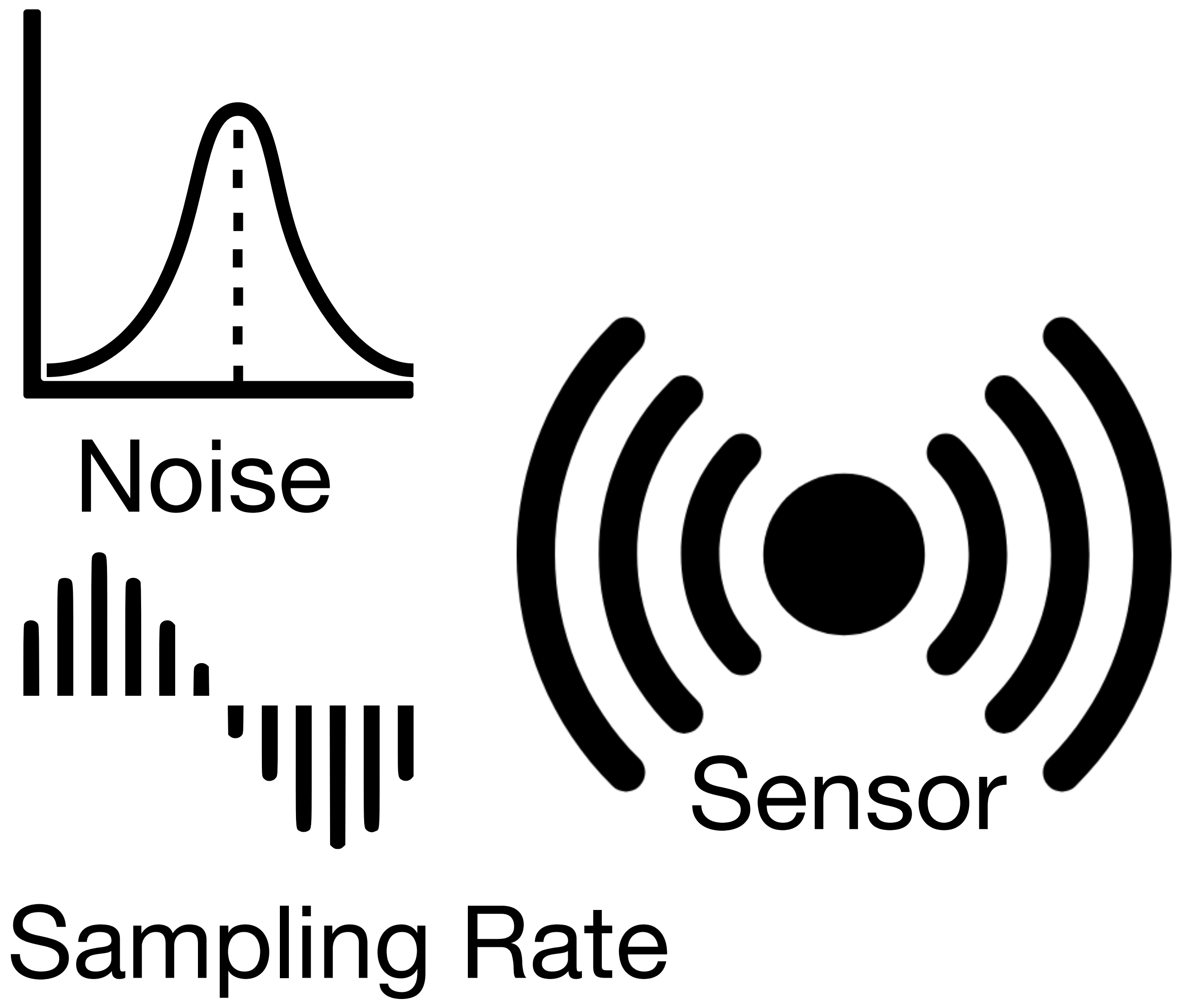

Noise
Sampling Rate
Sensor 1
Senses
Experiment 1
Noise
Sampling Rate
Sensor 2
Senses
Experiment 2
Noise
Sampling Rate
Sensor 3
Senses
Experiment 3
Noise
Sampling Rate
Sensor 4
Senses
Experiment 4
Random Series
Sinus
Peaks
Spikes
Square

This means we can not directly calibrate sensors

Because we do not know when they sensed the same



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

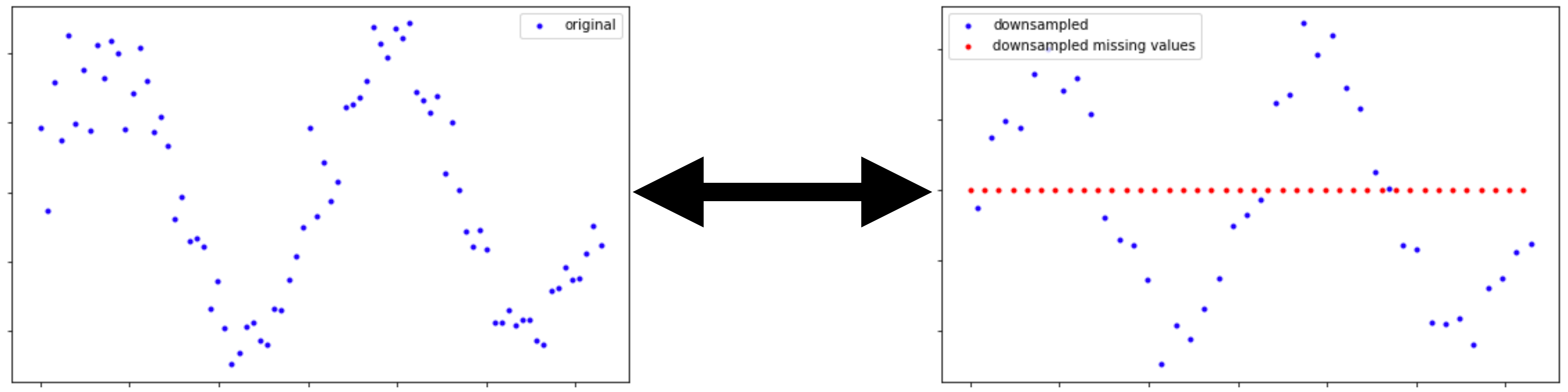


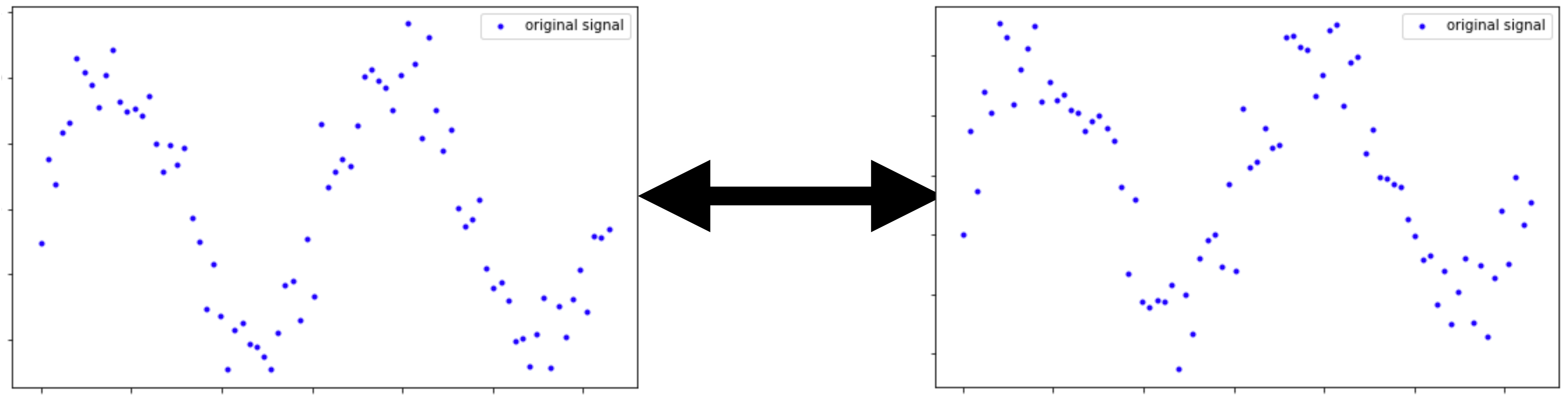
Common Representation (100Hz)

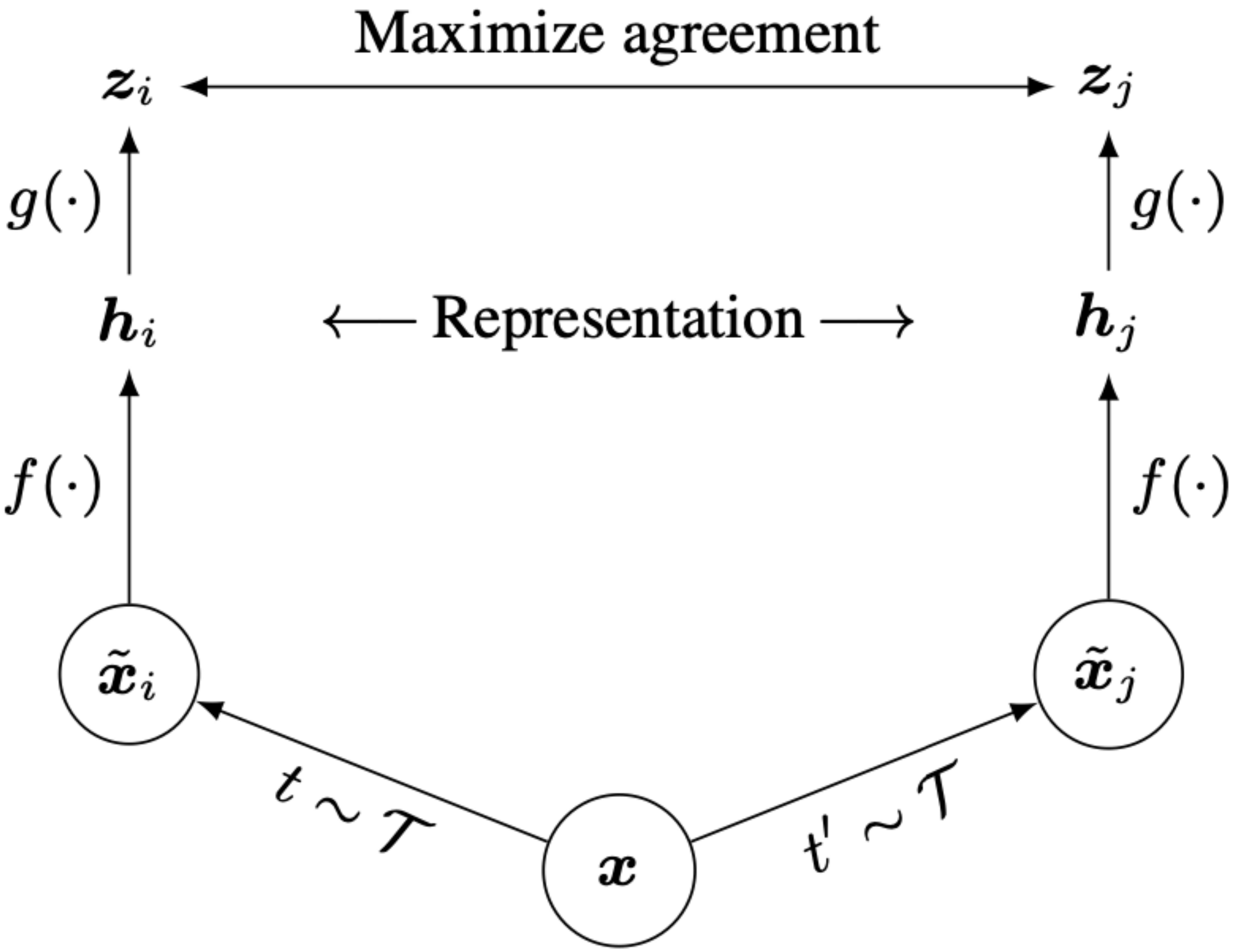
Noise

Sampling Rate

Ablations (sections of sensor failure)





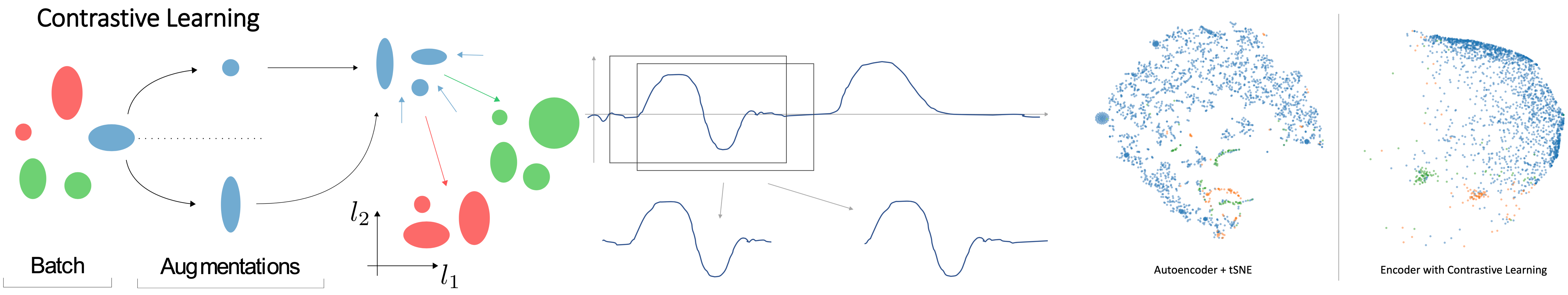

Maximize agreement
$\boldsymbol{z}_i$
$\boldsymbol{z}_j$
$g(\cdot)$
$g(\cdot)$
$\boldsymbol{h}_i$
$\longleftarrow$ Representation $\longrightarrow$
$\boldsymbol{h}_j$
$f(\cdot)$
$f(\cdot)$
$\tilde{\boldsymbol{x}}_i$
$\tilde{\boldsymbol{x}}_j$
$t \sim \mathcal{T}$
$t' \sim \mathcal{T}$
$\boldsymbol{x}$

Demo Time

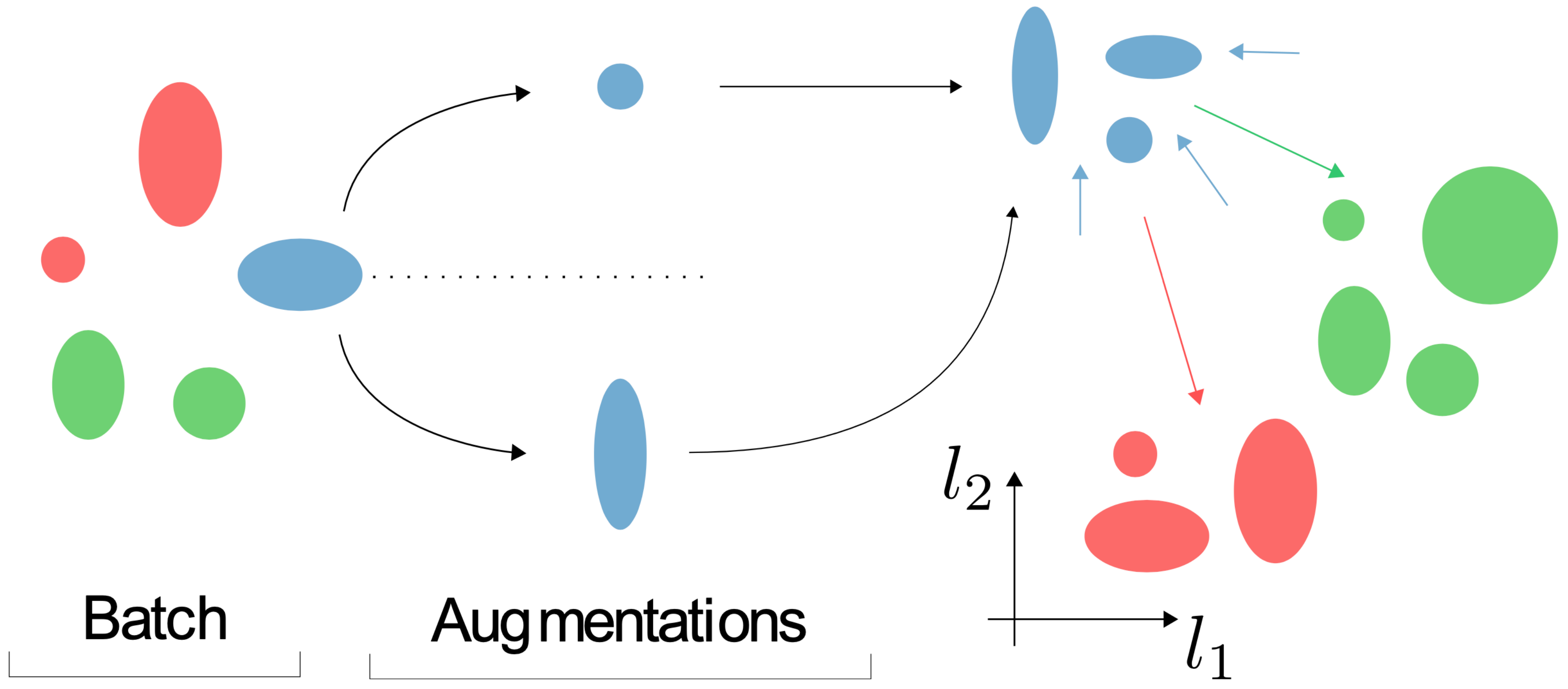
https://github.com/grazai/contrastive-learning-june-2020

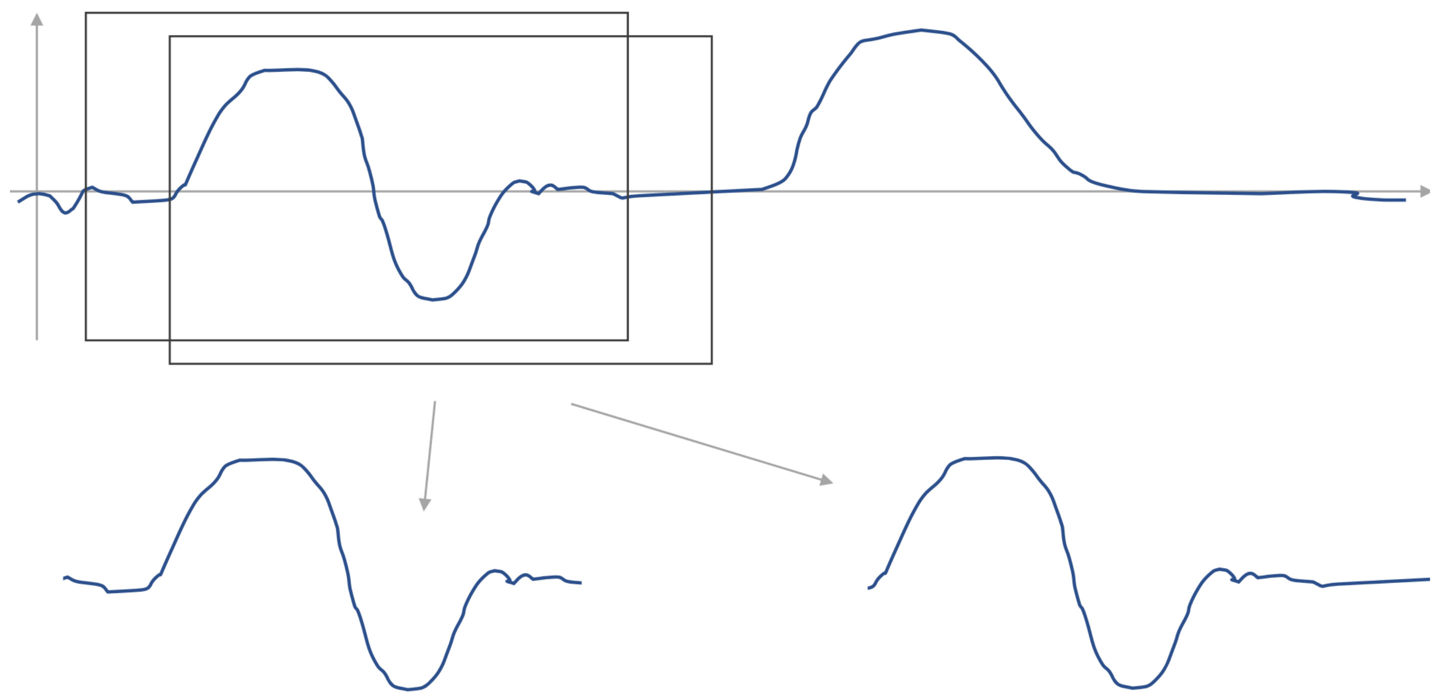
examples/different-sensors-jsimon

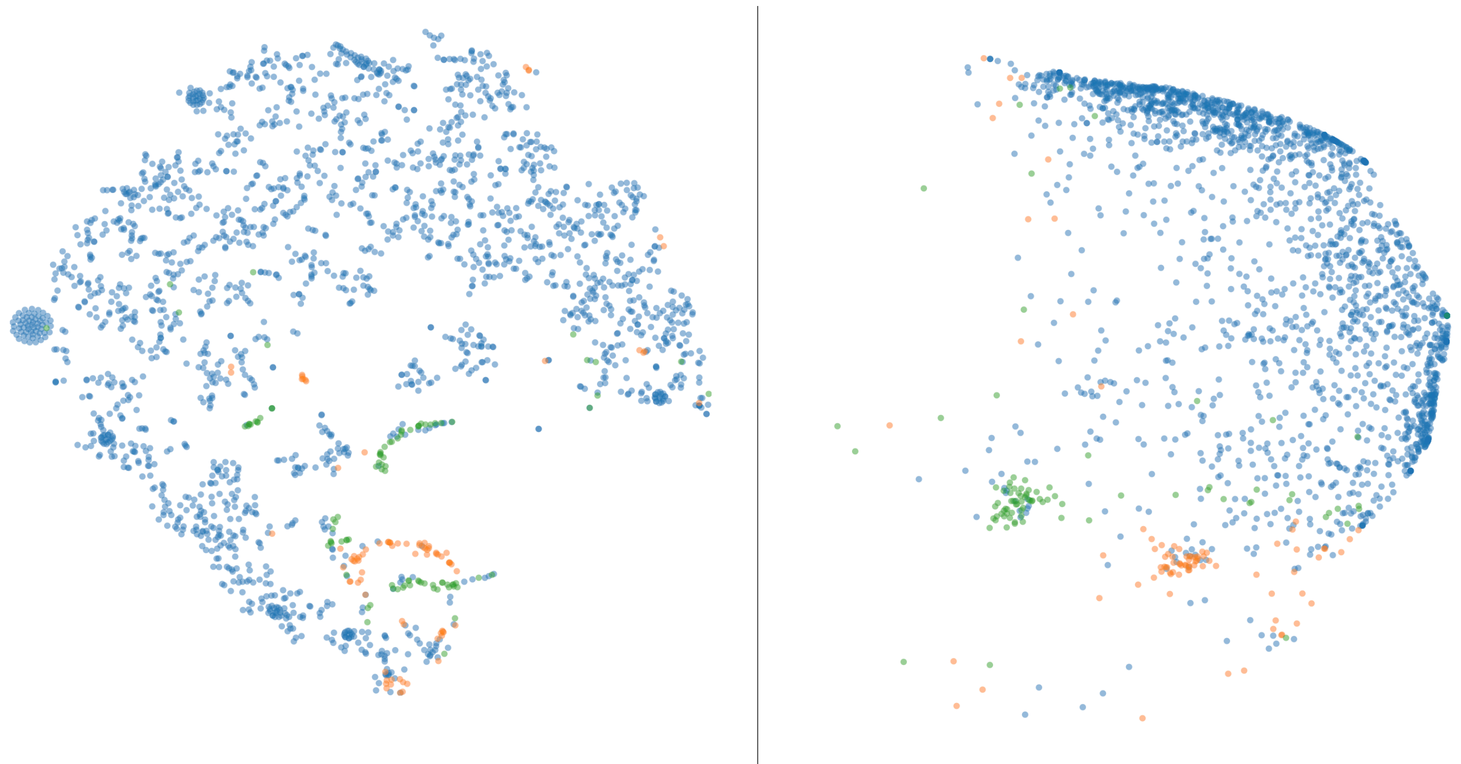
Summary


Contrastive Learning
Batch
Augmentations
l_2
l_1
Autoencoder + tSNE
Encoder with Contrastive Learning

Contrastive Learning



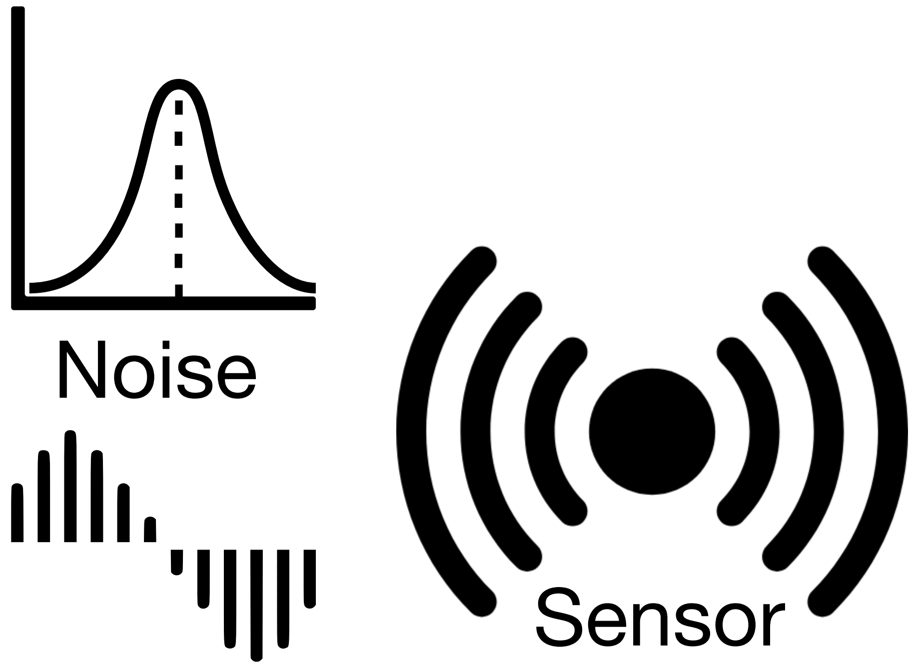


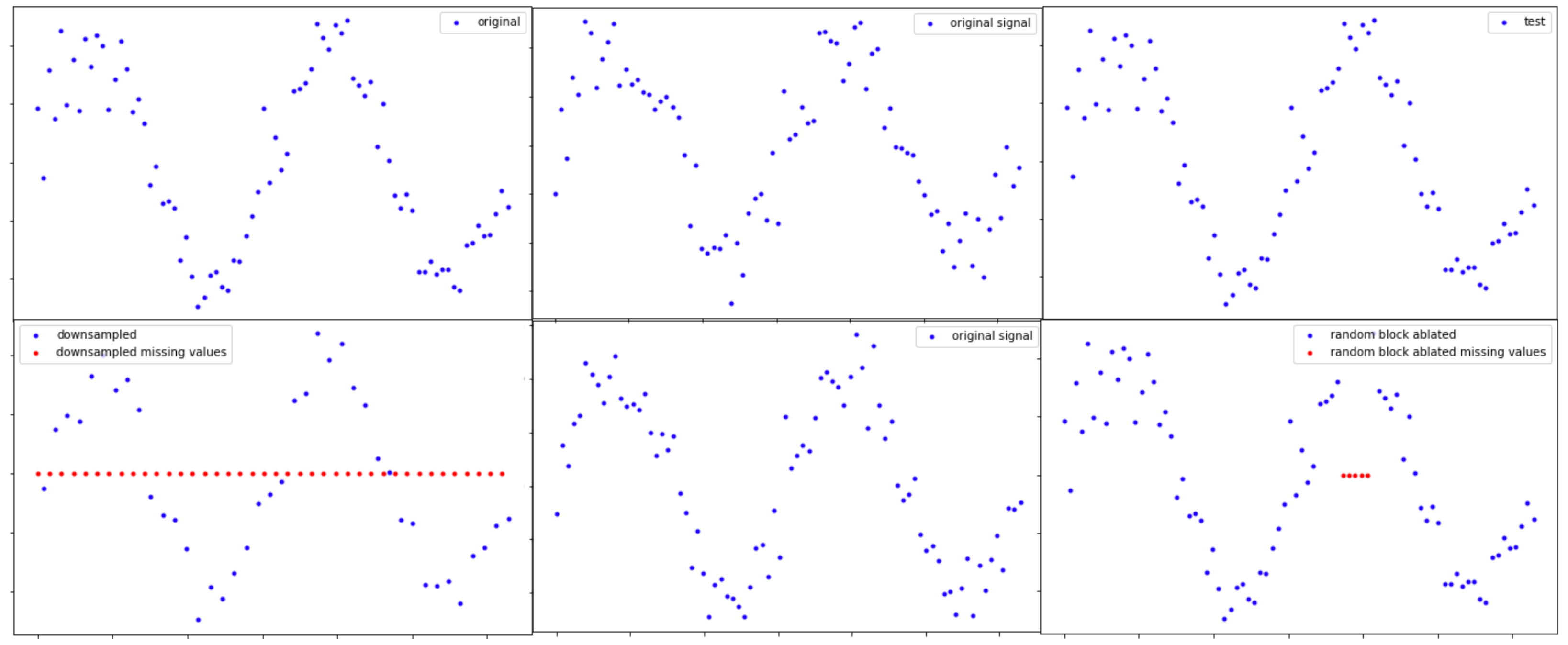


2005

2012

2020





Thanks for the Attention

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