

# Introduction to Representation learning: Approaches, Challenges and Applications

ABB



Shuyu Lin  
2016 AIMS cohort,  
Cyber Physical Systems Group,  
Department of Computer Science,  
University of Oxford

Vizle

This PDF is generated automatically by **Vizle**.  
Slides created *only for a few minutes* of your Video.



For the full PDF, please **Login to Vizle**.

<https://vizle.offnote.co> (Login via Google, top-right)

**Stay connected** with us:

Join us on **Facebook, Discord, Quora, Telegram**.



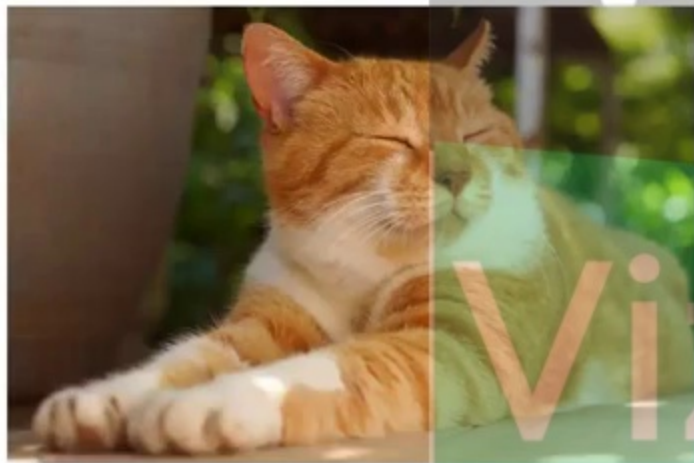
# What is Representation Learning?



# What is Representation Learning?



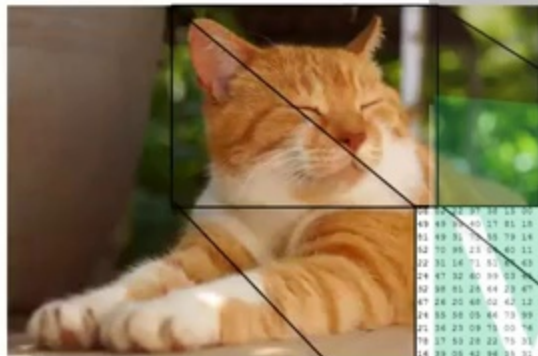
- What is the representation of a cat?



Vizle

# What is Representation Learning?

- Representation is a summary of data, which
  - Omits the unnecessary details,
  - And preserves important content.



69	89	91	20	27	81	18	37	40	37	17	40	31	43	89	40	04	86	42	00
81	89	81	76	55	79	14	29	83	71	40	87	93	48	30	53	69	13	34	48
62	70	88	21	40	11	42	49	14	40	34	31	32	54	71	37	32	34	91	
22	31	14	71	51	43	89	43	82	94	94	22	40	40	28	44	13	33	80	
29	47	32	80	39	03	02	84	79	33	53	70	38	84	29	15	17	12	30	
32	38	81	28	44	23	47	14	36	40	67	59	54	70	44	18	38	44	71	
47	24	20	68	02	42	12	20	18	94	39	40	08	40	91	44	49	98	21	
24	86	06	06	84	79	99	24	97	70	78	96	83	14	86	34	89	63	71	
21	34	23	09	75	00	76	44	21	93	14	00	41	33	97	34	32	33	39	
78	17	83	28	22	79	31	47	15	94	00	94	42	16	14	09	19	16	92	
14	39	00	42	94	31	32	47	89	88	00	24	17	94	24	34	29	88	37	
94	54	00	88	36	71	09	07	86	44	44	07	47	21	59	11	14	37	84	
19	80	81	49	06	94	47	49	28	79	92	13	64	02	17	04	09	51	40	
04	52	98	83	97	38	30	14	97	97	97	32	14	44	21	19	33	27	86	44
89	34	49	97	87	42	20	72	03	44	33	67	44	51	12	32	30	93	53	49
36	42	14	79	38	25	39	11	24	99	72	10	09	44	29	32	02	74	34	
20	88	34	41	72	30	23	81	34	42	89	49	82	47	89	85	14	24	14	
20	79	85	19	78	31	90	01	74	81	89	71	48	84	81	14	23	57	14	
61	70	84	71	83	51	84	49	14	82	33	49	41	43	52	01	89	19	47	

vizle

# Evolution of Representation Learning

Shuyu Lin



# Evolution of Representation Learning (continued)

End-to-end training of NN on task

Task  
↑  
Deep Model

Learning transferable features

Task 1 Task 2 Task 3

Features

Learnt from data

Deep Model

Good

Ideal



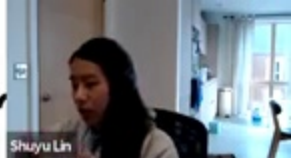
or



or

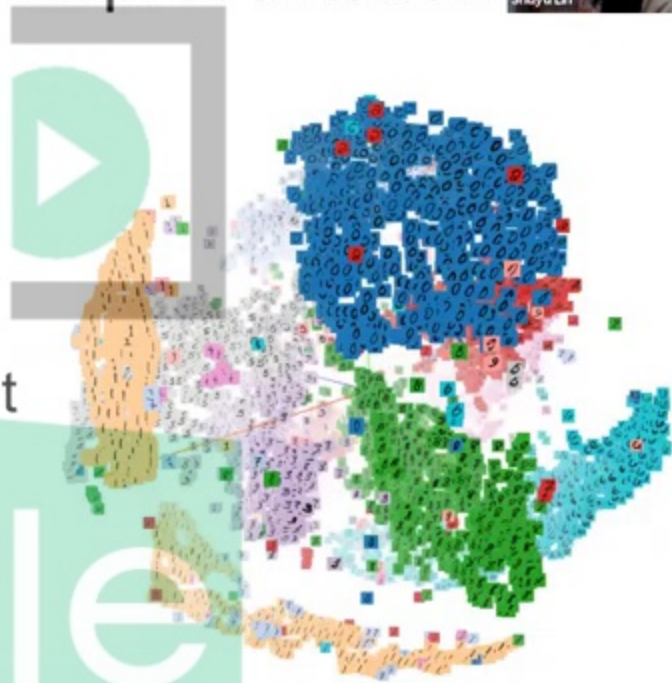


# What makes a Good Representation?



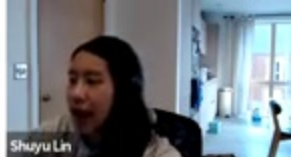
- Low dimensional
- Reusable across tasks
- Smooth and spatially coherent
- Disentangled
- Hierarchical and meaningful

Vizle

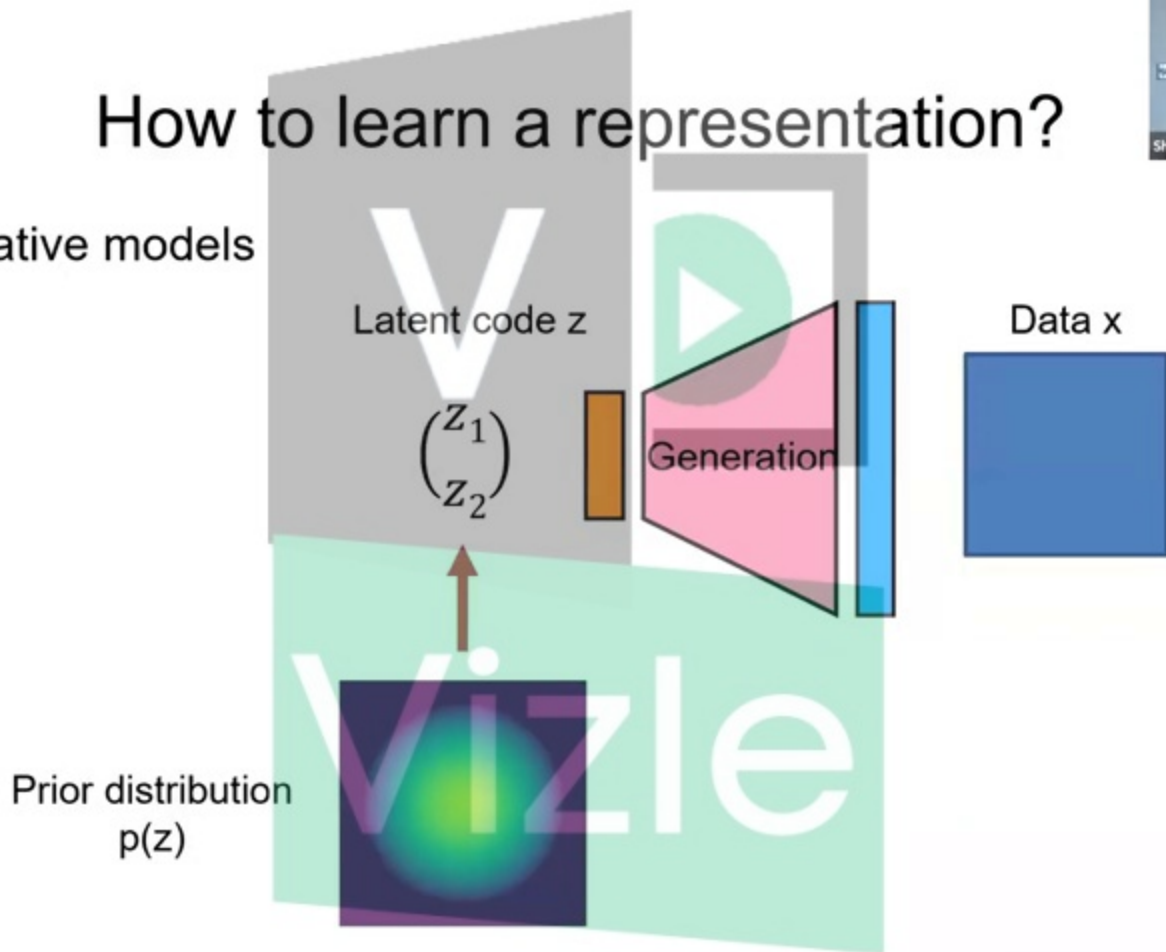




# How to learn a representation?

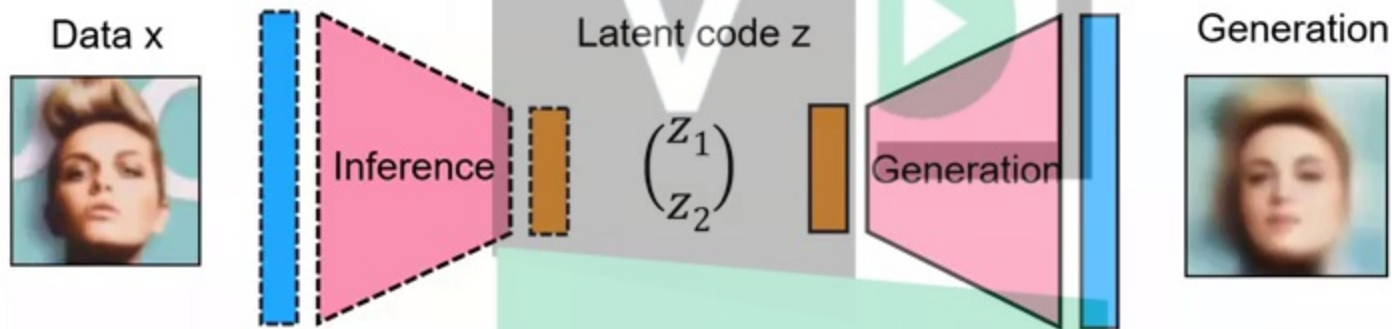


- Generative models



# How to learn a representation?

- Variational autoencoders (VAEs)



- What's the difference between VAEs and autoencoders?
  - Probabilistic formulation

# How to use a learnt representation?

- Anomaly detection in time series
- Latent space interpolation

 Vizle

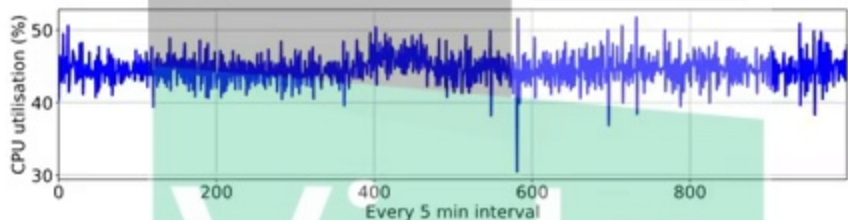
# Anomaly detection in time series

Example 1:  
Microsoft Share  
Price



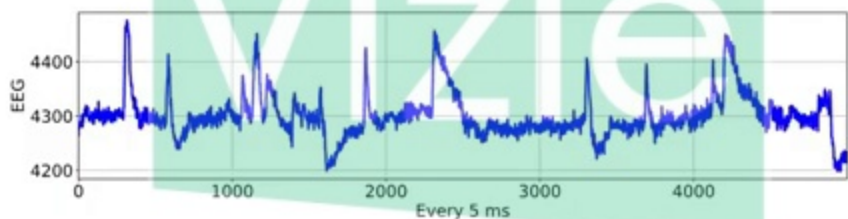
Avoid trading as  
the market is  
unusual?

Example 2:  
Amazon Server  
Utilization



Is there any attack  
to my service?

Example 3:  
A Patient EEG  
readings



Is my patient in  
a critical condition  
and needs  
intensive care?

# Anomaly detection is very difficult

Because:

1. Expertise and domain knowledge is often required to correctly identify an anomaly.



Labels are rare and expensive.

2. Anomalous events are rare and behaviours of anomalous events vary significantly from time to time.



Positive events are rare and hard to generalize.


As a result, anomaly detection for time series is practically an unsupervised learning task.

# A high-level sketch for anomaly detection

To respect the unsupervised learning nature of an anomaly detection task, we

Step 1: Aim to model the normal behaviours very well during training.

Training data: no anomaly labels.

Vizle

# How to use learnt representation for anomaly detection

Anomaly detection with a variational autoencoder (VAE) <sup>1</sup>





# Anomaly detection with VAE-LSTM hybrid model

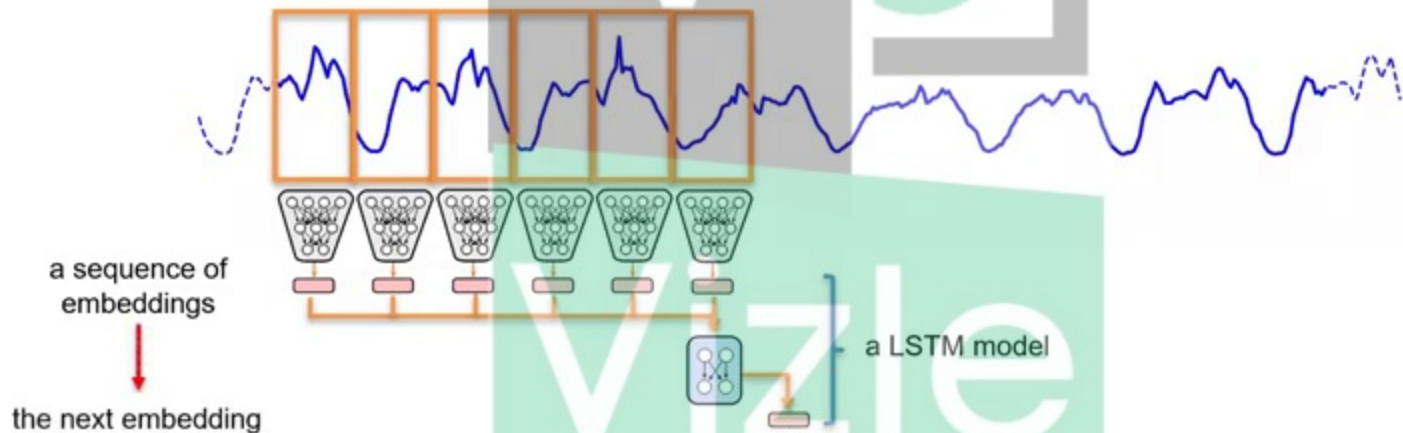
First, we train a VAE model to extract local information of a short window into a low-dimensional embedding.





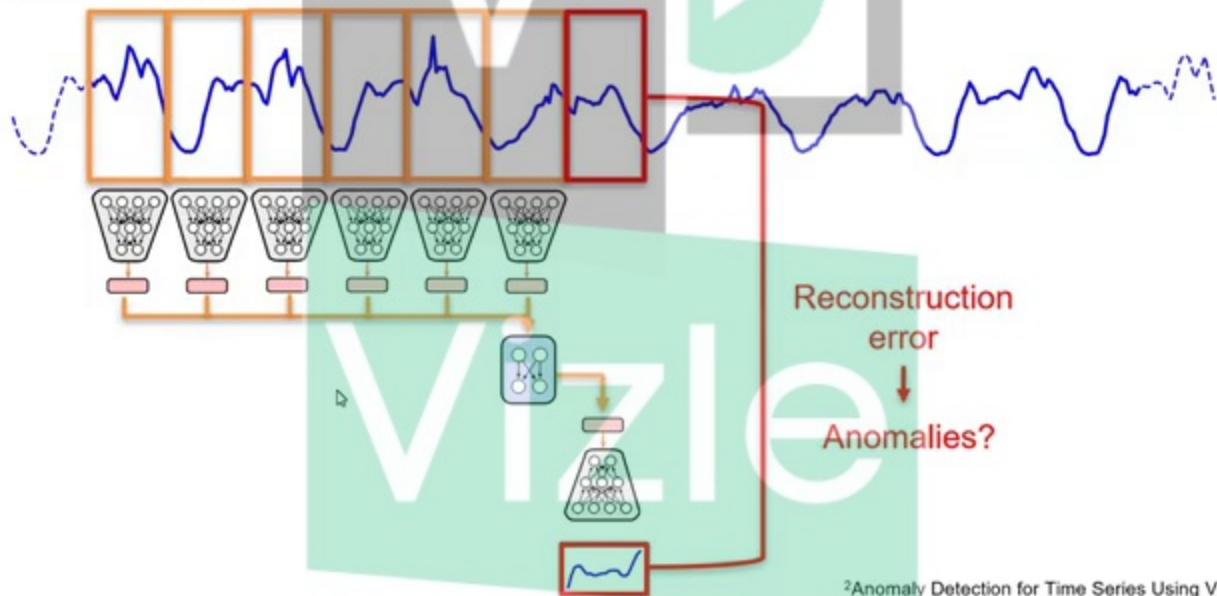
# Anomaly detection with VAE-LSTM hybrid model

Then, we use a LSTM model, which acts on the low-dimensional embeddings produced by the VAE model, to manage the sequential patterns over longer term.



# Anomaly detection with VAE-LSTM hybrid model

After both the VAE and the LSTM model are trained, we can detect anomalies using the reconstruction error of the next window.



# Anomaly detection with VAE-LSTM hybrid model

The benefits of our design are clear:

- Our VAE-LSTM hybrid model can extract information beyond a short local window.
- The VAE module forms **robust local features**.
- The LSTM module estimates the **long-term correlations** in the sequence.
- As a result, our detection algorithm can identify anomalies that might span over **multiple time scales**.



# Detection results



Our VAE-LSTM  
hybrid model

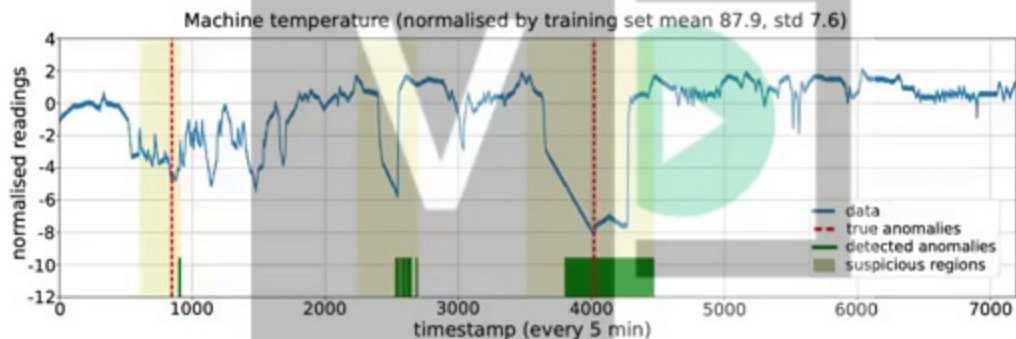
- Detect all anomalies
- 1 false positive, but it is a case worth further inspection

One false positive.  
Good for safety-critical situations.



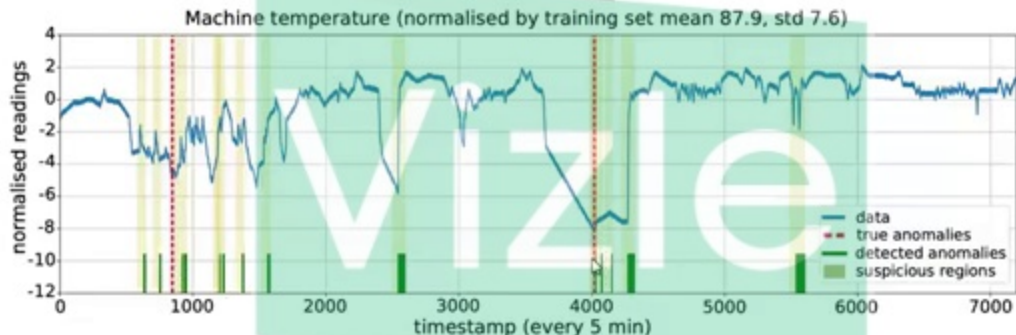
# Detection results

Our VAE-LSTM hybrid model



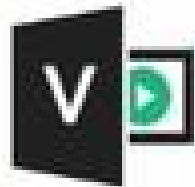
- Detect all anomalies
- 1 false positive, but it is a case worth further inspection

VAE model



- Detect all anomalies
- Too many false positives

This PDF is generated automatically by **Vizle**.  
Slides created *only for a few minutes* of your Video.



For the full PDF, please **Login to Vizle**.

<https://vizle.offnote.co> (Login via Google, top-right)

**Stay connected** with us:

Join us on **Facebook, Discord, Quora, Telegram**.