Tensor Network for Supervised Learning at Finite Temperature

Haoxiang Lin[1] Shuqian Ye[1] and Xi Zhu [1*]

[1]Shenzhen Institute of Artificial Intelligence and Robotics for Society
the Chinese University of Hong Kong, Shenzhen

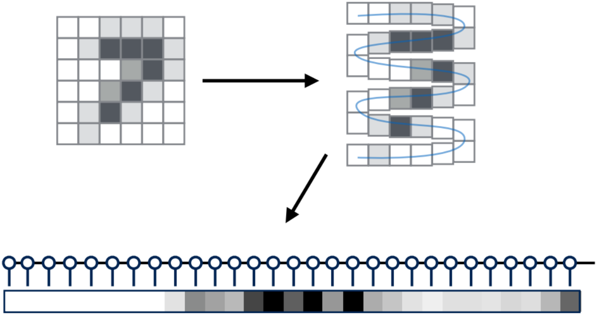
International Joint Conferences on Artificial Intelligence
workshop on Tensor Network representation on Machine Learning,
January 2021

Table of Contents

Table of Contents

Matrix Product State (MPS) classifier

Map image to the feature map through zigzag order.



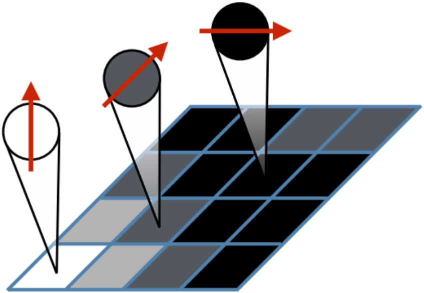
Feature map is the Kronecker product of local feature maps[1].

$$\mathbf{\Psi}(\boldsymbol{X}) = \mathbf{\Psi}^{S_1 S_2 \dots S_N}(\boldsymbol{p}) = \psi^{S_1}(p_1) \otimes \psi^{S_2}(p_2) \otimes \dots \psi^{S_N}(p_N)$$

[1] Edwin Stoudenmire and David J Schwab. Supervised learning with tensor networks. In Advances in Neural Information Processing Systems, pages 4799–4807, 2016.

Matrix Product State (MPS) classifier

Transform grayscale value $x \in [0, 1]$ into a local feature vector ψ.

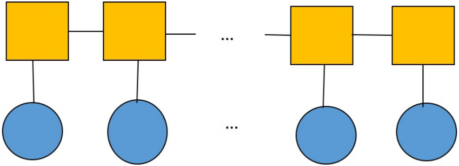


Example mapping:

$$\psi(x) = [\cos(\frac{\pi}{2}x), \sin(\frac{\pi}{2}x)]; \qquad \psi(x) = [x, 1 - x]$$

Matrix Product State (MPS) classifier

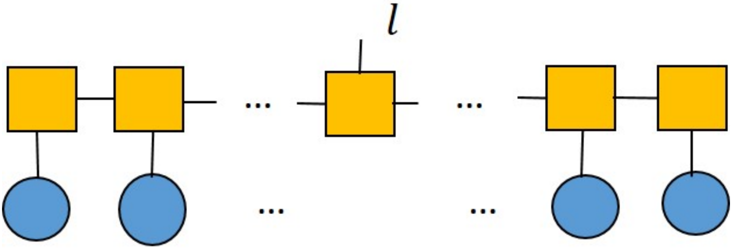
Yellow Cubic: the Matrix Product State (MPS)



Blue circle: the feature map.

Matrix Product State (MPS) classifier

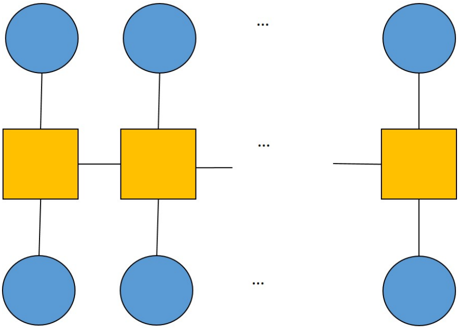
Yellow Cubic: the Matrix Product State (MPS)



Blue circle: the feature map.
For classification task, add an extra label tensor.

Minimally Entangled Typical Quantum States (METTS)

Yellow Cubic: the Matrix Product State (MPS), observable A in physics.



Blue circle: the feature map, wavefunction ψ in physics.
The contraction of it gives the observable $\langle\psi|A|\psi\rangle$.

Minimally Entangled Typical Quantum States (METTS)

If we consider the temperature effect[2]:

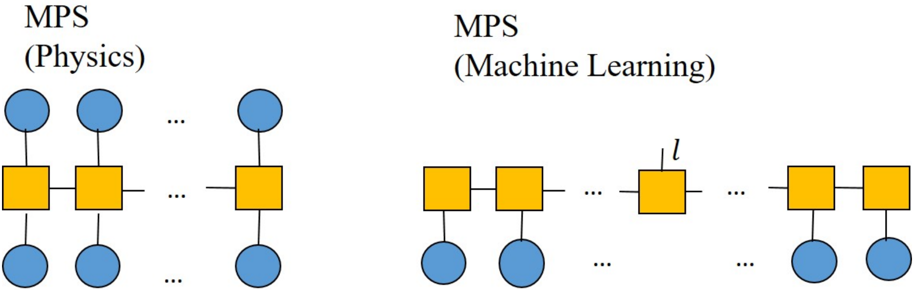
$$\langle A \rangle = \frac{1}{Z} \sum_i |\langle i e^{-\beta H/2} A e^{-\beta H/2} | i \rangle$$

In machine learning task,
Treat $|i\rangle$ as image
Treat A (MPS) as energy (H)

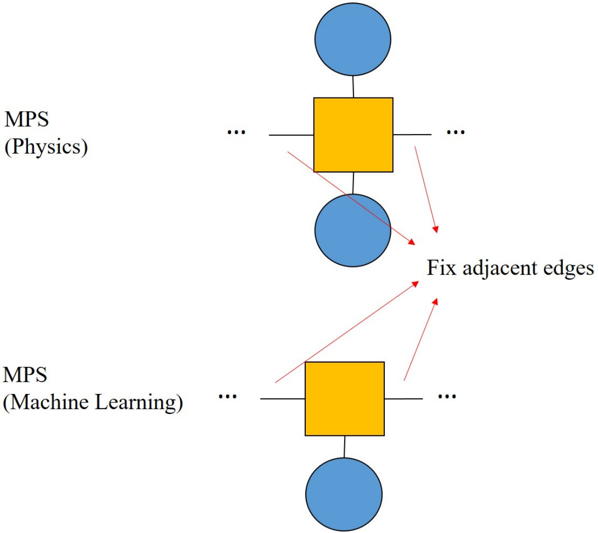
[2] EM Stoudenmire and Steven R White. Minimally entangled typical thermal state algorithms. New Journal of Physics, 12(5):055026, 2010.

Table of Contents

Machine Learning to Physics



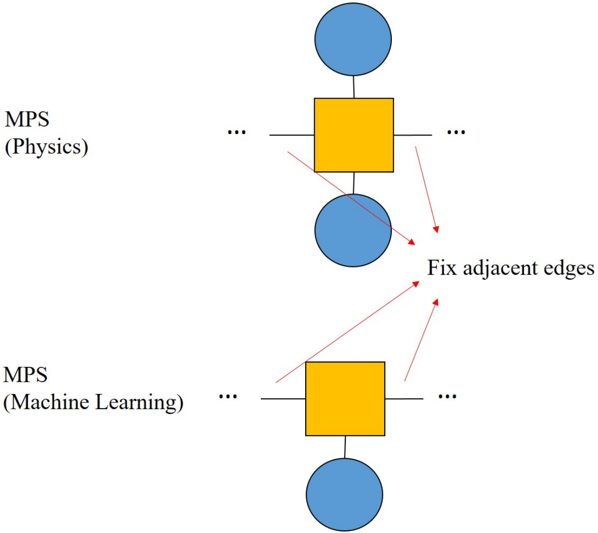
Machine Learning to Physics



$$A'[:,i,j,:] = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix}$$

$$A[:,i,:] = [A_1 \quad \ldots \quad A_n]$$

Machine Learning to Physics

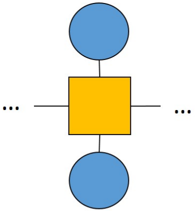


$$A'[:, i, j, :] = \begin{bmatrix} A_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_n \end{bmatrix}$$

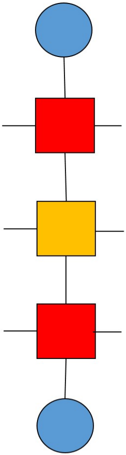
diagonal element

$$A[:, i, :] = [A_1 \quad \ldots \quad A_n]$$

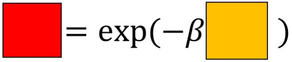
Insertion of Temperature Layer



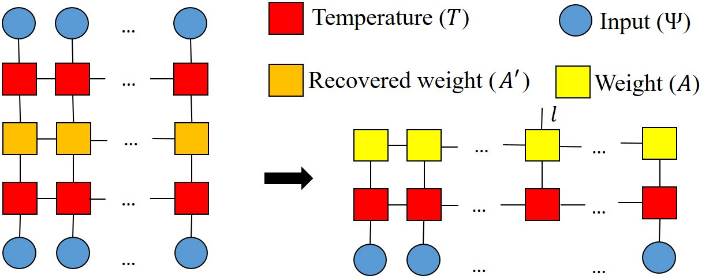
MPS
without temperature



MPS
With temperature



Insertion of Temperature Layer

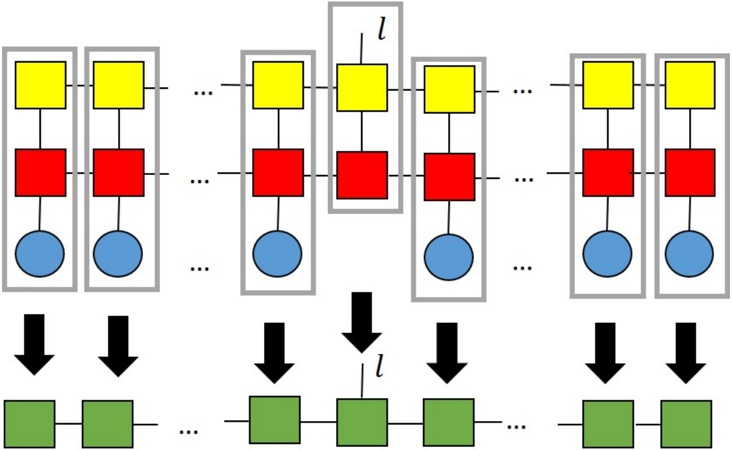


Until now the Finite Temperature Tensor Network (FTTN) has constructed.

Table of Contents

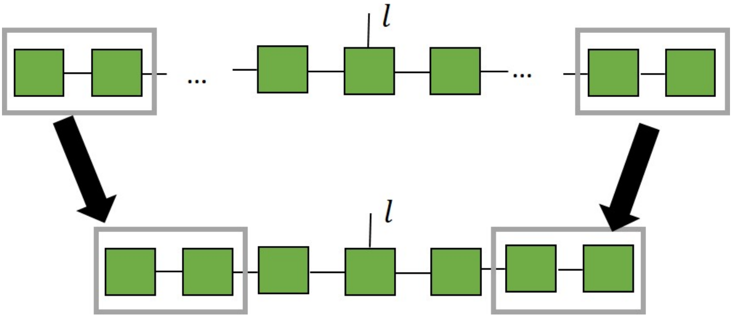
Parallel Contraction Algorithm

Step 1:



Parallel Contraction Algorithm

Step 2: Contract in pairs.



Step 3: repeat step 2 until converge.

Table of Contents

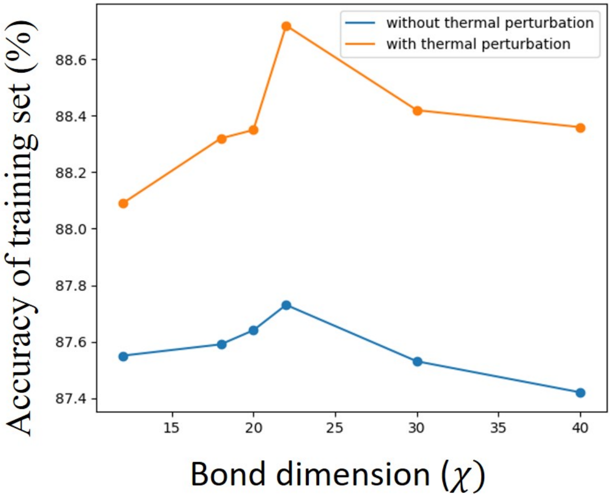
Same setup as [3]

Dataset: Fashion-MNIST

- Optimizer: Adam
- Learning Rate: 1e-4
- Batch Size: 50
- Image Size: 28×28
- Local Feature Map: $\psi(x) = [x, 1 - x]^T$
- Loss function: multi-class cross-entropy $\text{Loss} = \frac{1}{2} \sum_{n=1}^{N_T} \sum_l (f^l(\boldsymbol{x}_n) - \boldsymbol{y}_n^l)$

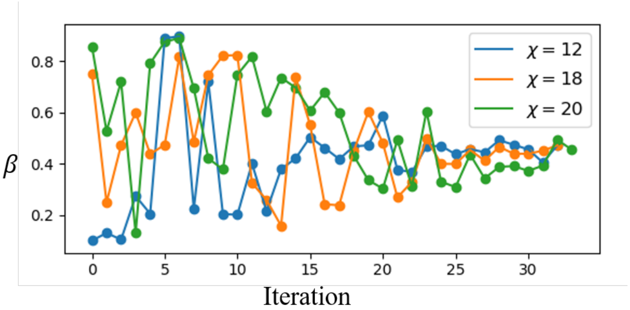
[3] Stavros Efthymiou, Jack Hidary, and Stefan Leichenauer. Tensornetwork for machine learning. arXiv preprint arXiv:1906.06329, 2019.

Experiment results



Interesting Discovery

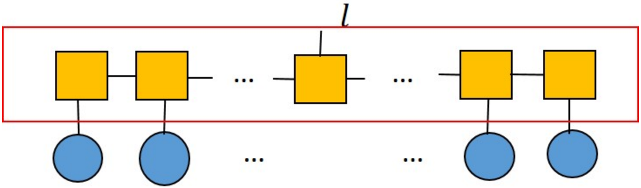
We tried to optimize temperature-like parameter β by simulated annealing algorithm.



This parameter is nearly independent of bond dimension χ.

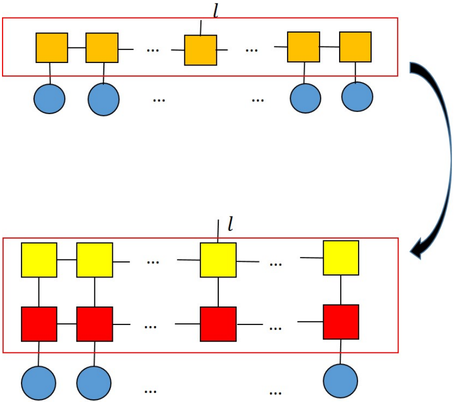
Table of Contents

Physical Interpretation



MPS can also represent a feature map $|\rho\rangle$.
Contraction gives inner product, the result comes from the largest one.

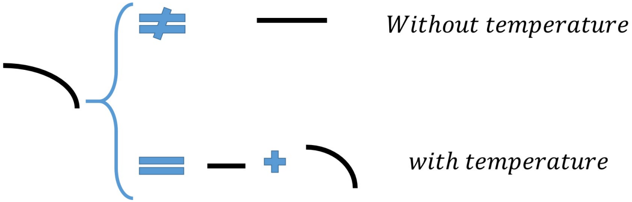
Physical Interpretation



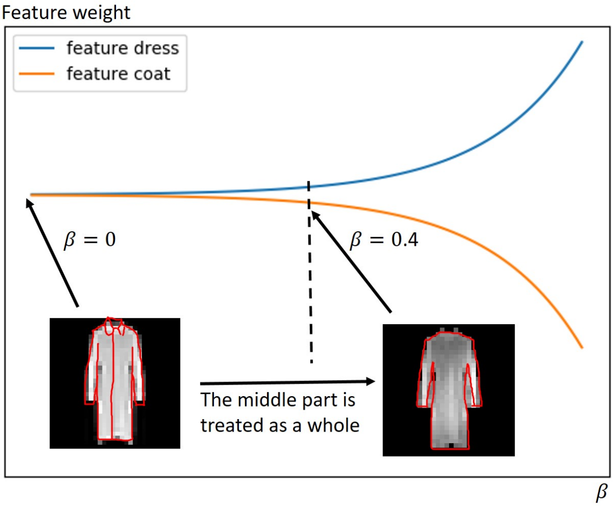
$|\rho\rangle.$

$e^{-\beta H}|\rho\rangle.$

Physical Interpretation



Physical Interpretation

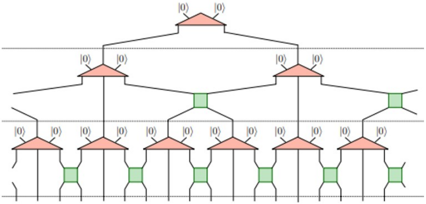


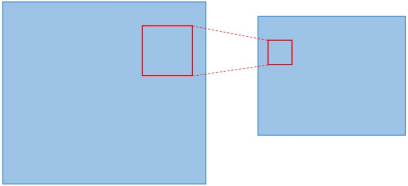
Outlook

Multi-scale Entangled Renormalization Ansatz (MERA)

Similar structure

Convolutional Neural Network





References

[1] **MPS classifier:** Edwin Stoudenmire and David J Schwab. Supervised learning with tensor networks. In Advances in Neural Information Processing Systems, pages 4799–4807, 2016.

[2] **METTS algorithm:** EM Stoudenmire and Steven R White. Minimally entangled typical thermal state algorithms. New Journal of Physics, 12(5):055026, 2010.

Thanks for listening