Tensor Network for Supervised Learning at Finite Temperature

Haoxiang Lin\textsuperscript{1}  Shuqian Ye\textsuperscript{1}  and  Xi Zhu \textsuperscript{1*}

\textsuperscript{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>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction: MPS classifier and METTS algorithm</td>
</tr>
<tr>
<td>2</td>
<td>Architecture of FTTN: the insertion of temperature layer</td>
</tr>
<tr>
<td>3</td>
<td>Contraction and Optimization Algorithm</td>
</tr>
<tr>
<td>4</td>
<td>Experiment Result and Interesting Discovery</td>
</tr>
<tr>
<td>5</td>
<td>Physical Interpretation and Outlook</td>
</tr>
</tbody>
</table>
Matrix Product State (MPS) classifier

Map image to the feature map through zigzag order.

Feature map is the Kronecker product of local feature maps\(^1\).

\[
\Psi(X) = \Psi^{S_1S_2\ldots S_N}(p) = \psi^{S_1}(p_1) \otimes \psi^{S_2}(p_2) \otimes \ldots \psi^{S_N}(p_N)
\]

Matrix Product State (MPS) classifier

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

Example mapping:

$$\psi(x) = [\cos\left(\frac{\pi}{2}x\right), \sin\left(\frac{\pi}{2}x\right)]; \quad \psi(x) = [x, 1 - x]$$
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 $\psi$ in physics. The contraction of it gives the observable $\langle \psi | A | \psi \rangle$. 
If we consider the temperature effect:

\[
\langle A \rangle = \frac{1}{Z} \sum_i \langle ie^{-\beta H/2} A e^{-\beta H/2} i \rangle
\]

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

---

### Table of Contents

1. Introduction: MPS classifier and METTS algorithm

2. Architecture of FTTN: the insertion of temperature layer

3. Contraction and Optimization Algorithm

4. Experiment Result and Interesting Discovery

5. Physical Interpretation and Outlook
Finite Temperature Tensor Network

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

\[ A[i, : , :] = [A_1 \ldots A_n] \]
Machine Learning to Physics

MPS (Physics)

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

Fix adjacent edges

diagonal element

MPS (Machine Learning)

\[ A[:, i, :] = [A_1 \ldots A_n] \]
Insertion of Temperature Layer

MPS without temperature

\[ \text{red} = \exp(-\beta \text{yellow}) \]

MPS with temperature
Insertion of Temperature Layer

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

Until now the Finite Temperature Tensor Network (FTTN) has constructed.
# Table of Contents

1. Introduction: MPS classifier and METTS algorithm

2. Architecture of FTTN: the insertion of temperature layer

3. Contraction and Optimization Algorithm

4. Experiment Result and Interesting Discovery

5. Physical Interpretation and Outlook
Parallel Contraction Algorithm

Step 1:
Step 2: Contract in pairs.

Step 3: repeat step 2 until converge.
1. Introduction: MPS classifier and METTS algorithm
2. Architecture of FTTN: the insertion of temperature layer
3. Contraction and Optimization Algorithm
4. Experiment Result and Interesting Discovery
5. Physical Interpretation and Outlook
Same setup as $^3$

**Dataset: Fashion-MNIST**
- Optimizer: Adam
- Learning Rate: 1e-4
- Batch Size: 50
- Image Size: $28 \times 28$
- Local Feature Map: 
  $$\psi(x) = [x, 1 - x]^T$$
- Loss function: 
  multi-class cross-entropy
  $$\text{Loss} = \frac{1}{2} \sum_{n=1}^{N_T} \sum_l (f_l(x_n) - y_n^l)$$

---

Experiment results

![Graph showing accuracy of training set (%) vs. bond dimension (χ)]

- **Accuracy of training set (%)**
- **Bond dimension (χ)**

**Legend:**
- Blue line: without thermal perturbation
- Orange line: with thermal perturbation

---

H. Lin, S. Ye, X. Zhu (AIRS)

Finite Temperature Tensor Network

IJCAI 2020 TNML
Interesting Discovery

We tried to optimize temperature-like parameter $\beta$ by simulated annealing algorithm.

This parameter is nearly independent of bond dimension $\chi$. 

![Graph showing the behavior of $\beta$ over iterations for different $\chi$ values.](image)
1. Introduction: MPS classifier and METTS algorithm

2. Architecture of FTTN: the insertion of temperature layer

3. Contraction and Optimization Algorithm

4. Experiment Result and Interesting Discovery

5. Physical Interpretation and Outlook
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

Without temperature

with temperature
Physical Interpretation

Feature weight

- Blue line: feature dress
- Orange line: feature coat

\[ \beta = 0 \]

\[ \beta = 0.4 \]

The middle part is treated as a whole

\[ \beta \]
Outlook

Multi-scale Entangled Renormalization Ansatz (MERA)

Similar structure

Convolutional Neural Network

Thanks for listening