

# Progress of works

Taoyuan Wang

2017/12/19

# Pytorch

- Word Prediction

# Word Embedding

- Word Vector
  - 12-word
  - 7-dimension

# Training Data

- Pairs

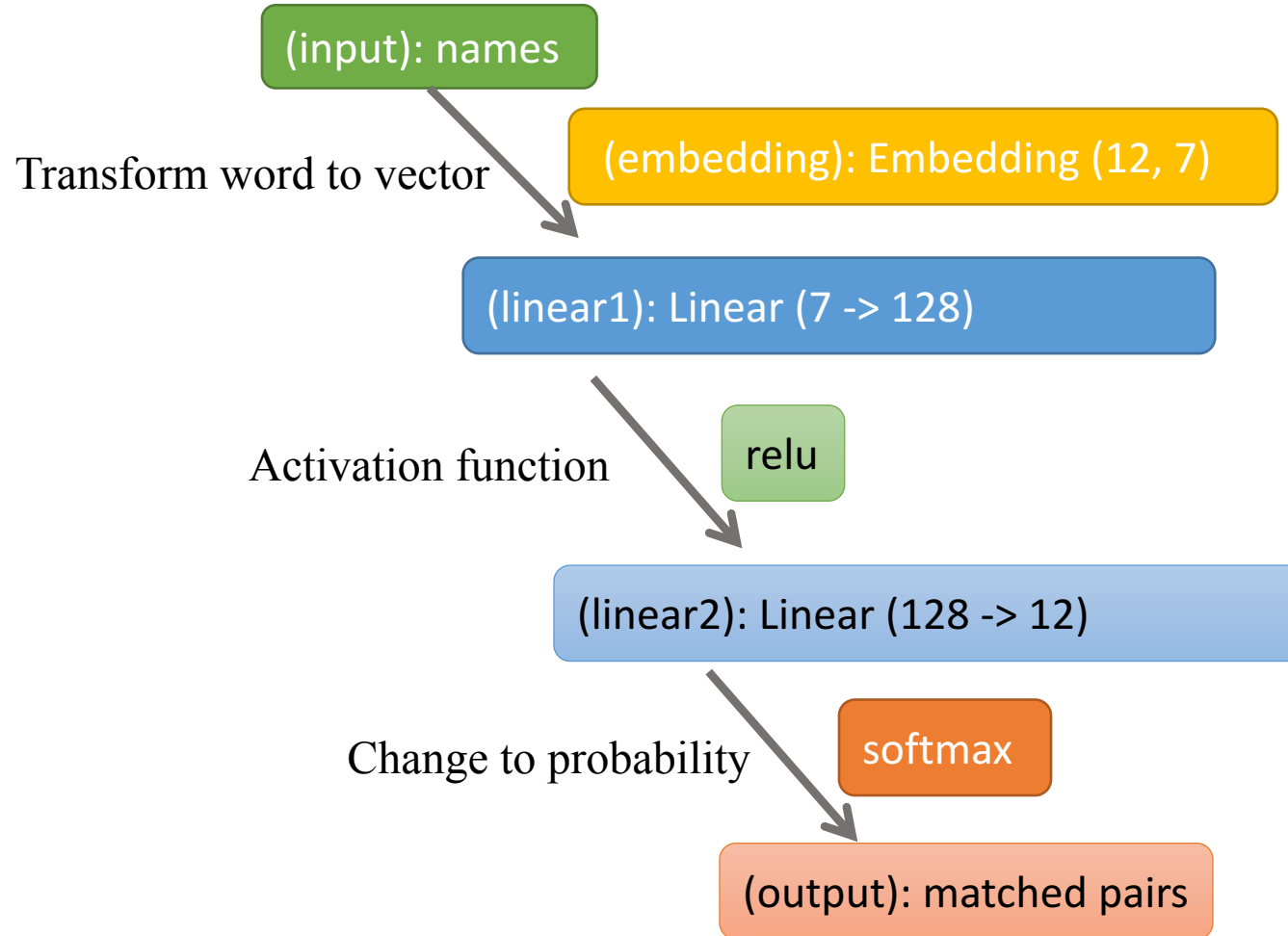
word	label
Noah	Liam
Mason	Jacob
William	Ethan
James	Alexander
Michael	Benjamin
abcd	dcba

# Word Indexing

- Dictionaries

0: Noah	1: Liam
2: Mason	3: Jacob
4: William	5: Ethan
6: James	7: Alexander
8: Michael	9: Benjamin
10: abcd	11: dcba

# Bi-gram Model



# Training

- Loss Function: CrossEntropyLoss
- Optimizer: SGD
  - Learning Rate: 1e-3
- Epoch: 100

# Testing

```
Loss: 0.630986
epoch: 96
*****
Loss: 0.626964
epoch: 97
*****
Loss: 0.622974
epoch: 98
*****
Loss: 0.619016
epoch: 99
*****
Loss: 0.615089
epoch: 100
*****
Loss: 0.611193

word is Noah, predict is Liam, lable is Liam
word is Mason, predict is Jacob, lable is Jacob
word is William, predict is Ethan, lable is Ethan
word is James, predict is Alexander, lable is Alexander
word is Michael, predict is Benjamin, lable is Benjamin
word is abcd, predict is dcba, lable is dcba
```

word	label
Noah	Liam
Mason	Jacob
William	Ethan
James	Alexander
Michael	Benjamin
abcd	dcba



# Future Work

- Sentence Data
  - Cloze Test
- Other Models
  - LSTM
  - Attention
  - Autoencoder
- CUDA

# Chatbot

- Trace STC code and rebuild one
- Improve it