



Experience Sharing Report

Mieradilijiang Maimaiti
20210608



Outline

- About me
- Past of my life
 - Before Ph.D.
- Study life
- Research life
- Daily life
- Job experiences
- Miscellaneous

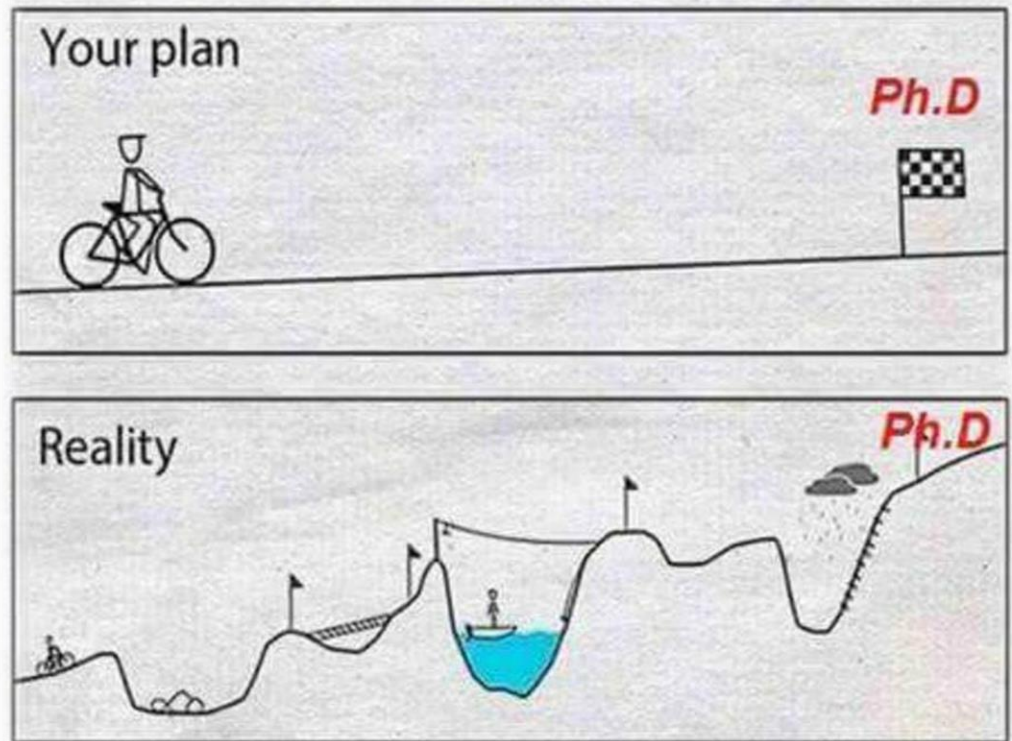


Photo from google





About me

- Original name
 - Meanings
- Nick name
 - How was it originated from
- Hometown
- Educational background
- Family members
 - My parents
 - Younger brother



About me

- Original name
 - Mieradilijiang Maimaiti
 - 米尔阿迪力江·麦麦提
 - Name: 米尔+阿迪力+江
 - Prefix: “米尔” represents **adjectives**
 - Stem: “阿迪力” represents **fair; fairness**
 - Suffix: “江” represents **warm; respect**
- Family name
 - Muhammad: 穆罕穆德
 - Maimaiti: 麦麦提



About me

- Nick name
 - 木拉丁
 - When I was in junior middle school I was name by one of my Chinese friends
 - Murad : 木拉提
 - Muradil: 木拉丁
 - Younger (<32): 木哥; 木师兄; 木拉丁
 - Elder (>32): 木拉丁; 小木;
 - Social network friends: 木木





About me

- Hometown
- Kashgar, Xinjiang



Photos from google



About me

- Education background



B.S., Computer Science and Technology

M.S., Software Engineering



Ph.D. Computer Science and Technology



About me

- Family members





Past of my life : bachelor

- Before my PhD life
 - Bachelor
 - Software engineer
 - Desktop application
 - Web application
 - Smart mobile application
 - Android + WP
 - Projects
 - Win7
 - Linux
 - Open office
 - Hand writing identification
 - C++





Past of my life : bachelor

Top navigation bar: (H) ياردەم (I) ئىقتىدار (C) ئېقىتساد باشقۇرۇش (W) كۆزەك (S) ئىزدەش (U) تەھرىرلەش (L) ئۇچۇر باشقۇرۇش

Left sidebar (Top):

- ئوقۇتقۇچى قوشۇش
- ئوقۇغۇچى قوشۇش
- ئوقۇغۇچى ئوقۇتقۇچى قوشۇش
- دەرس سىنىپ قوشۇش
- مۇناش بىكىتىش
- ئىزدەش
- ھەق باشقۇرۇش
- يوقلما قىلىش
- كېرىم ئەھۋالى

Left sidebar (Bottom):

- ئوقۇتقۇچى قوشۇش
- دەرس سىنىپ قوشۇش
- مۇناش بىكىتىش
- ئىزدەش
- ھەق باشقۇرۇش
- يوقلما قىلىش
- كېرىم ئەھۋالى

Main Content Area (Top):

- تېلېفون نومۇرى: 0995-2525255
- ئۆچ ئايلىق: 13039558574
- كىملىك نومۇرى: 653101198504010688
- ئىمەيل: fdsfsdfs@126.com
- ئىزدەش: 00:09:09 2012-05-20
- كەلگەن ۋاقىت: 09:09 2012-02-20
- تەللىغان دەرس: 1 خەنزۇ تىلى ئاساس 1-قىسىم
- تەللىغان دەرس: 2 ئىنگلىز تىلى ئېغىز تىلى كۈچەيتىش

Main Content Area (Bottom):

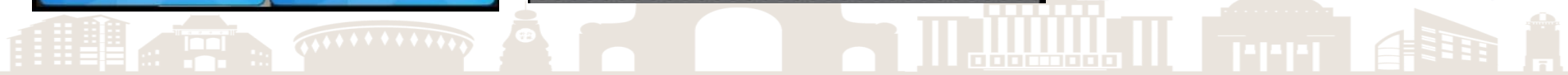
- ئىزدەش شەرتى: 25-03-2014
- تەللىغان دەرس: 25-03-2014
- تەللىغان دەرس: 25-03-2014
- تەللىغان دەرس: 25-03-2014
- تەللىغان دەرس: 25-03-2014

Table (Bottom):

ئىمەيل	كىملىك نومۇرى	ئۆچ ئايلىق	تېلېفون نومۇرى	جىنسى	ئوقۇغۇچى ئىسمى	نومۇرى
@126.com	653101198905010722	0991-8585856	13039448657	✓	مۇستاپا مىرزا	1
@126.com	653101198504010688	0995-2525255	13039558574	✓	بېشى	2
gdgdf@12	65320119890501058X	0991-2525256	13039448575	✓	تالغان قۇنى Excel غا چىقىرىش	3
@126.com	653101198905010433		13039558574	✓	ھەسسى چىقىرىش	4
	653101198905010544		13036559574	✓		5

Bottom status bar: 16:38:32 2014-يىلى 3-ئاينىڭ 25-كۈنى سائەت 16:20:05 كىرگەن ۋاقتى! hero كىردى! باشقۇرغۇچى: 16:38

Past of my life : bachelor



Past of my life : master

- Before my PhD life
 - Master
 - Software engineer
 - ASR
 - TTS
 - MT
 - Projects
 - Public analysis
 - POS-Tagging
 - Morphology
 - MT





Siri:

能听懂你话的手机语音助手





Past of my life : master





Past of my life : master

汉-维, 维-汉在线翻译系统 x

localhost/mosesweb/ Google

 新疆多语种信息技术实验室
شىنجاڭ كۆپ تىللىق ئۇچۇر تېخنىكا تەجرىبىخانىسى

输入

☐ 汉文 → 维文 ☒ 维文 → 汉文

هاۋا بەك سوغۇق، شۇبرغان چىقۇاتىدۇ

天气很冷，风来

清除文本

翻译 ☐ 查看词对齐



Past of my life : master



你好。

ياخشىمۇ سىز .



今天天气怎么样？

بۈگۈن ھاۋا قانداقراق ؟



今天天气很好。

بۈگۈن ھاۋا ناھايىتى ياخشى .



بۈگۈن ھاۋا ناھايىتى ياخشى .



ئەتە ھاۋا قانداق

明天天气怎么样



ھازىر سائەت نەچچە بولدى

现在几点了



ئەتە بىز مەكتەپكە بارىمىز

明天我们要去学校



明天去学校干什么？

ئەتە مەكتەپكە بېرىپ نېمە ؟





Past of my life : master



作者

米尔阿迪力江·麦麦提



结果中检索

高级检索

知识元检索 >

引文检索 >

总库

9

中文

外文

科技

社科

☐ 技术研究(1)

主题

发表年度

文献来源

学科

作者

机构

基金

文献类型

广告

中科院材料1区 7分国际刊

纳米光子学

高引TOP10 免费读

学术期刊

8

学位论文

1

会议

0

报纸

0

年鉴

图书

0

专利

标准

0

成果

0

检索范围: 总库 作者: 米尔阿迪力江·麦麦提

主题定制

检索历史

共找到 9 条结果

☐ 全选 已选: 0 清除

批量下载

导出与分析

排序: 相关性 发表时间↓ 被引 下载

显示 20



	题名	作者	来源	发表时间	数据库	被引	下载	操作
☐ 1	基于智能手机的维吾尔语语音控制系统的开发	米尔阿迪力江·麦麦提; 吾守尔·斯拉木; 努尔麦麦提·尤鲁瓦斯; 热依曼·吐尔逊; 艾尼宛尔·托乎提	计算机应用与软件	2016-06-15	期刊	3	104	
☐ 2	基于Morfessor的维吾尔语词干提取和词性标注的研究	米尔阿迪力江·麦麦提	新疆大学	2015-05-24	硕士	4	305	
☐ 3	基于.NET的维哈柯多语种网上数据采集系统的设计与实现	再吐娜木·阿巴白克力; 侯存义; 米尔阿迪力江·麦麦提; 张立新	电脑知识与技术	2015-04-15	期刊	1	94	
☐ 4	丝绸之路经济带相似语言信息横向处理通信技术的研究	伊·达瓦; 米尔阿迪力江·麦麦提	新疆师范大学学报(自然科学版)	2014-12-30	期刊	84		
☐ 5	基于Windows Mobile平台的产品管理软件的设计与开发	米尔阿迪力江·麦麦提; 吾守尔·斯拉木; 热依曼·吐尔逊; 迪丽尼格尔·热夏提	新疆大学学报(自然科学版)	2014-11-15	期刊	38		
☐ 6	基于智能设备的汉语学习软件的研究与开发	米尔阿迪力江·麦麦提; 吾守尔·斯拉木; 迪丽尼格尔·热夏提; 帕尔哈提·达吾提	新疆大学学报(自然科学版)	2014-05-15	期刊	7	202	
☐ 7	基于统计机器翻译技术的胡都木—托忒文本转写的实现	伊·达瓦; 王美慧; 米尔阿迪力江·麦麦提	西部蒙古论坛	2014-05-15	期刊	83		

Past of my life : master

- Before my PhD life
 - No idea about what is the science?
 - No idea about what does the research like?
 - Idea
 - Experiment
 - Writing
 - Analysis
 - No idea about what is the real group meeting?
 - It was just a report about the progress of the project
 - What should be discussed in the group meetings?
 - Lack of Face to face communication or discussion with mentor



Study life

- Almost from **scratch**

- NLP
- Machine Learning
 - Corresponding tools
 - scikit-learn, genism ...
- Deep Learning
 - Theory
 - Corresponding tools
 - Theano, TF, Pytorch
- Math, linear algebra
- English
- Python
 - Numpy, scipy

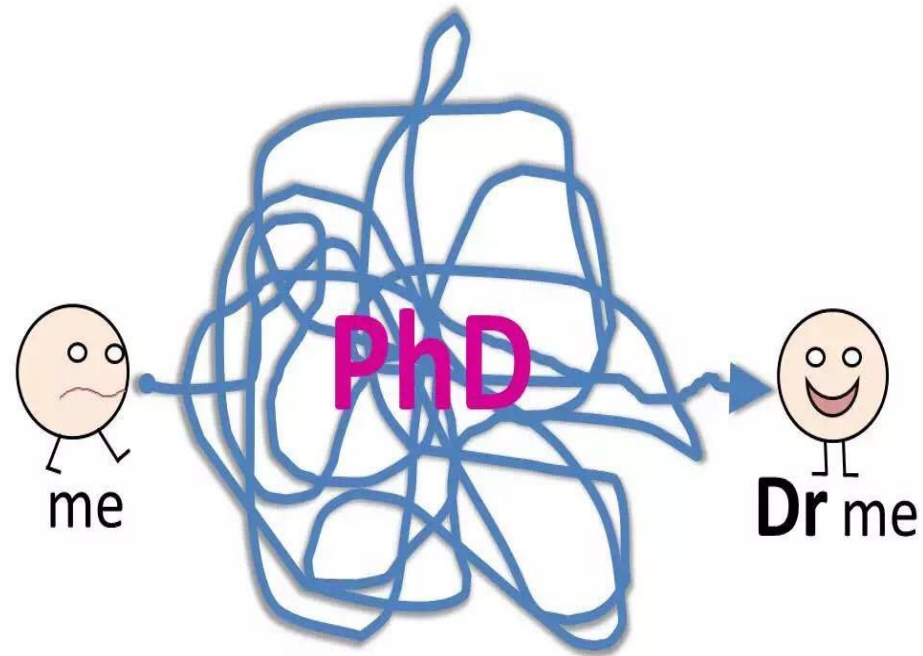


Photo from google



20

Study life: Math

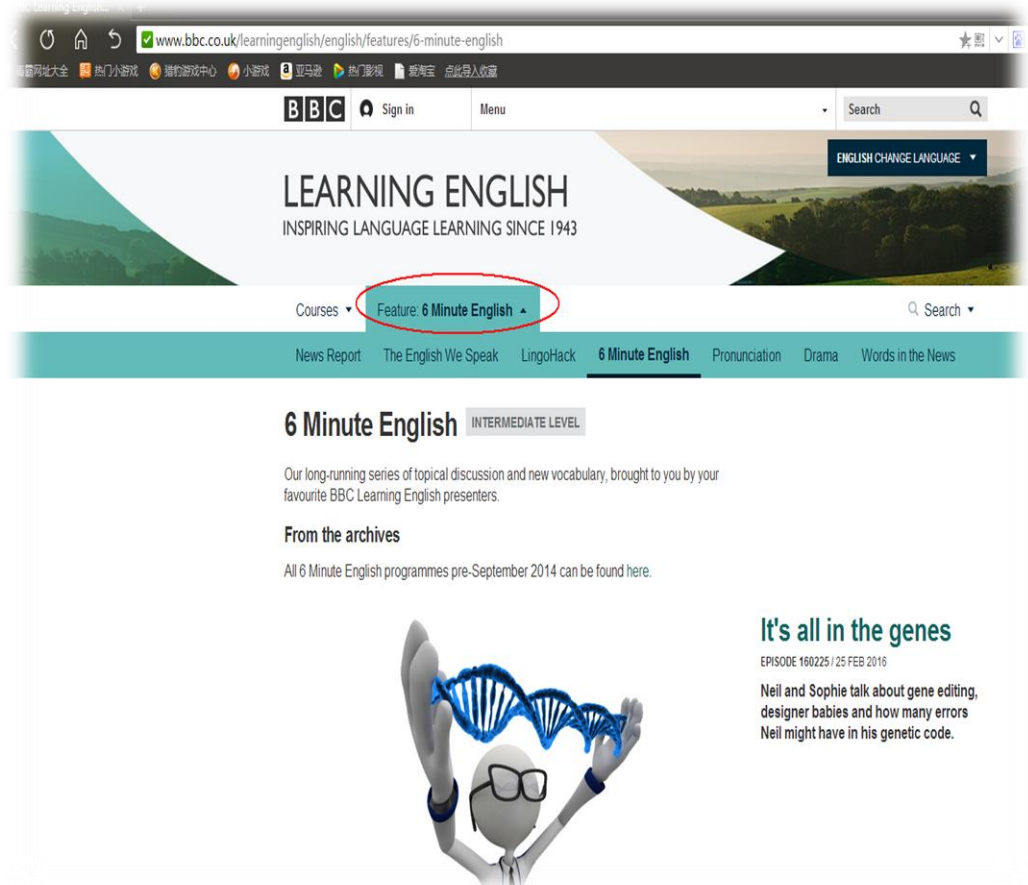
- Functions
- Function limit
- Bozalno
- Chauchy
- Function limit arithmetic's
-





Study life: English

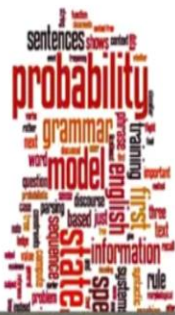
- Listening
- Vocabulary
- Speaking
- 每日英语
- (有道口语大师)



<http://www.bbc.co.uk/learningenglish/english/features/6-minute-english>



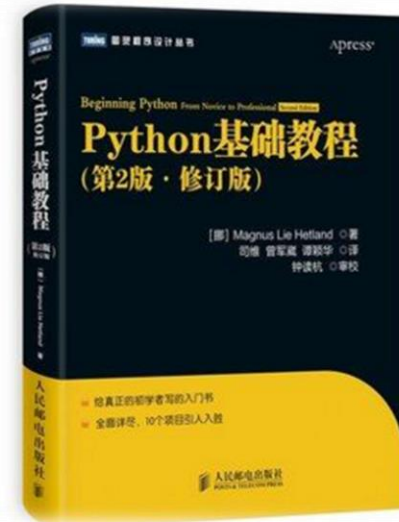


- 
- # Introduction to NLP
- What is Natural Language Processing?
- 



Study life: Python

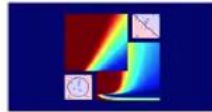
- Condition
- Range && xrange
- Function definition
- Recursive
- Permutation
- Variables area
- String methods
- Dictionary && methods
-



Study life: ML

- Applications of Machine Learning
- Components of Machine Learning
- Machine Learning and Other Fields
- Linear regression
- Logistic regression
- Some samples
-

Machine Learning Foundations
(機器學習基石)



Lecture 1: The Learning Problem

Hsuan-Tien Lin (林軒田)
htlin@csie.ntu.edu.tw

Department of Computer Science
& Information Engineering
National Taiwan University
(國立台灣大學資訊工程系)



Hsuan-Tien Lin (NTU CSIE) Machine Learning Foundations 0/27





Study life: DL

- RNN, LSTM, GRU
- BP
- CNN
- Dropout
- 3 layer NN
- NMT
 - Groundhog
- Tools
 - Theano
 - TF
 - Pytorch
 -

```
theano.tensor.patternbroadcast(x, broadcastable) \[source\]
```

Make the input adopt a specific broadcasting pattern.

```
theano.tensor.addbroadcast(x, *axes) \[source\]
```

Make the input broadcastable in the specified axes.

For example, `addbroadcast(x, 0)` will make the first dimension of `x` broadcastable. When performing the function, if the length of `x` along that dimension is not 1, a `ValueError` will be raised.

We apply the opt here not to pollute the graph especially during the gpu optimization

Parameters:

- `x` (*tensor_like*) – Input theano tensor.
- `axis` (*an int or an iterable object such as list or tuple of int values*) – The dimension along which the tensor `x` should be broadcastable. If the length of `x` along these dimensions is not 1, a `ValueError` will be raised.

Returns: A theano tensor, which is broadcastable along the specified dimensions.

Return type: `tensor`





Study life: DL

```
2016-09-29 00:23:16,371: groundhog.layers.basic: DEBUG: Loading W_0_dec_prev_readout_0 of (620, 1000)
Traceback (most recent call last):
  File "experiments/nmt/sample.py", line 311, in <module>
    main()
  File "experiments/nmt/sample.py", line 243, in main
    beam_search.compile()
  File "experiments/nmt/sample.py", line 44, in compile
    self.comp_next_probs = self.enc_dec.create_next_probs_computer()
  File "/data/disk1/private/mld/umda4nmt_ch_en_1/experiments/nmt/encdec.py", line 1605, in create_next_probs_computer
    self.c, self.step_num, self.gen_y, self.current_states)],
  File "/data/disk1/private/mld/umda4nmt_ch_en_1/experiments/nmt/encdec.py", line 1385, in build_next_probs_predictor
    given_init_states=init_states, step_num=step_num)
  File "/data/disk1/private/mld/umda4nmt_ch_en_1/experiments/nmt/encdec.py", line 1225, in build_decoder
    hidden_result = self.mul(hidden_layers[0], get_class[:, :, 0, :])
  File "/data/disk1/private/mld/umda4nmt_ch_en_1/groundhog/layers/basic.py", line 378, in __getitem__
    new_obj.out = new_obj.out.__getitem__(pos)
  File "/usr/local/lib/python2.7/dist-packages/theano/tensor/var.py", line 532, in __getitem__
    lambda entry: isinstance(entry, Variable)))
  File "/usr/local/lib/python2.7/dist-packages/theano/gof/op.py", line 611, in __call__
    node = self.make_node(*inputs, **kwargs)
  File "/usr/local/lib/python2.7/dist-packages/theano/tensor/subtensor.py", line 484, in make_node
    raise exception
ValueError: The index list is longer (size 4) than the number of dimensions of the tensor(namely 2). You are asking for a dimension of the tensor that does not exist! You might need to use dimshuffle to add extra dimension to your tensor.
```



Study life: DL

- 3 layer neural network
 - Basic fundamentals
 - Pure Numpy code

神经网络入门

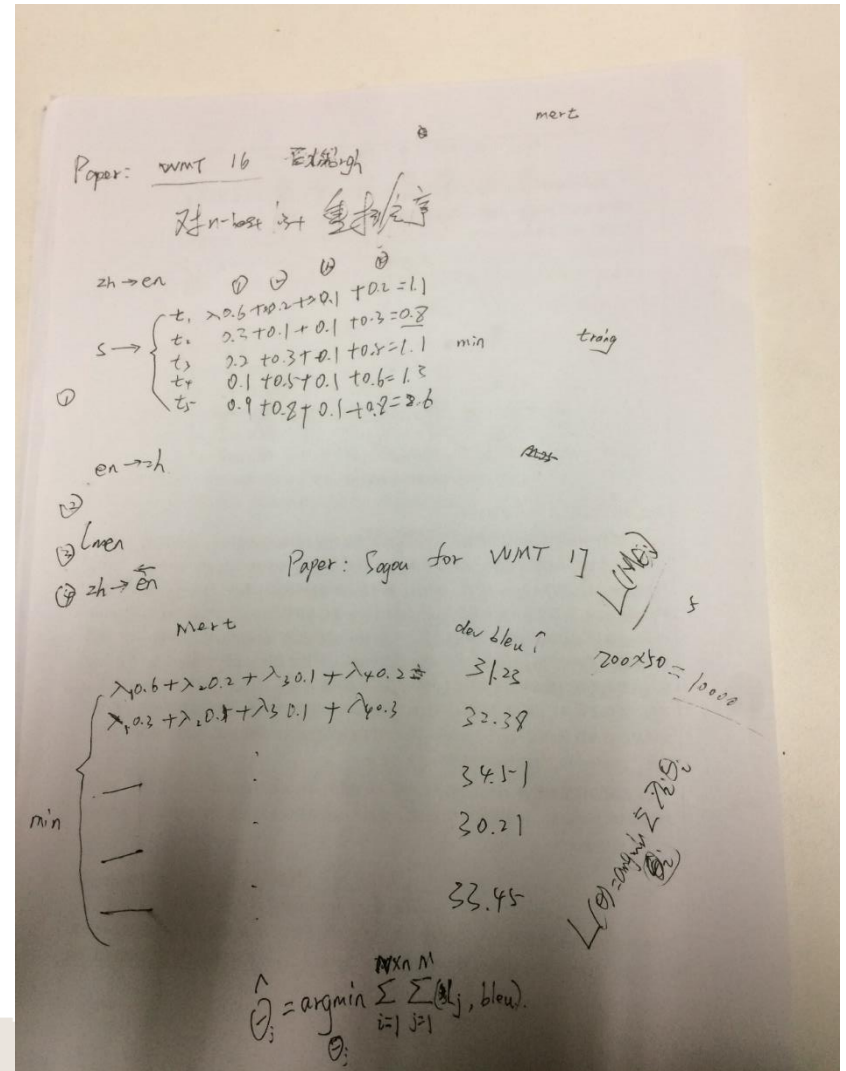
2015 年 12 月 8 日

```
解决方案.txt x second_competition_corpus_url.txt x my_mnist.py x
1 #following code is an original three layer neural_net
2 #which has not used any other deeplearning tool kit such as
3 #theano,caffee,tensorflow,torch,pytex and pytorch etc.
4
5 #import numpy as np
6 import numpy
7 #import cPickle as pickle
8 import cPickle
9 import itertools
10 import pdb
11
12 class Tanh:
13
14     def __init__(self):
15         self.error = None
16         self.output = None
17         self.gradient = []
18         self.parameter = []
19
20     def compute_output(self,x):
21         self.output = numpy.tanh(x) #numpy has predefined the activation function "tanh"
22
23     def compute_error(self,x,e):
24         # _ represents extra parameter and 'e' stands for errors achieved from previous layer
25         self.error = e * (1-self.out ** 2) #mainly calculate the derivate
26
27     def forward(self,x):
28         self.compute_output(x)
29         return self.output
30     def backward(self,x,e):
31         self.compute_error(x,e)
32         return self.error
33
34 class Sigmoid:
35
36 class Softmax:
37
38 #activation function (non-linear) has no any parameters, and gradients does not need to be calcul
39
40 class Linear:
41
42 class NLL:
43
44 #apply both the softmax and negative log likelihood
45
46 class SoftmaxNLL:
47
48 # 3-layer feed-forward neural network
49 class FNN:
50
51     def sgd(learning_rate, params, grads):
52
53     def rundataset(model, x,y,batch):
54
55     def main():
56
57 if __name__ == "__main__":
58     main()
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
```



Study life: DL

```
train.py  sample.py  encdec.py  score.py
25 from groundhog.models import LM_Model
26 from groundhog.datasets import PytablesBitextIterator
27 from groundhog.utils import sample_zeros, sample_weights_orth, init_bias
28 import groundhog.utils as utils
29
30 logger = logging.getLogger(__name__)
31
32 def create_padded_batch(state, x, y, return_dict=False): ==
131
132 def get_batch_iterator(state): ==
187
188 class RecurrentLayerWithSearch(Layer): ==
500
501 class ReplicateLayer(Layer): ==
515
516 class PadLayer(Layer): ==
529
530 class ZeroLayer(Layer): ==
535
536 def none_if_zero(x): ==
540
541 class Maxout(object): ==
559
560 def prefix_lookup(state, p, s): ==
564
565 class EncoderDecoderBase(object): ==
658
659 class Encoder(EncoderDecoderBase): ==
802
803 class Decoder(EncoderDecoderBase): ==
1262
1263 class RNNEncoderDecoder(object): ==
1493
1494 def parse_input(state, word2idx, line, raise_unk=False, idx2word=None, l
1517
```



N-best list

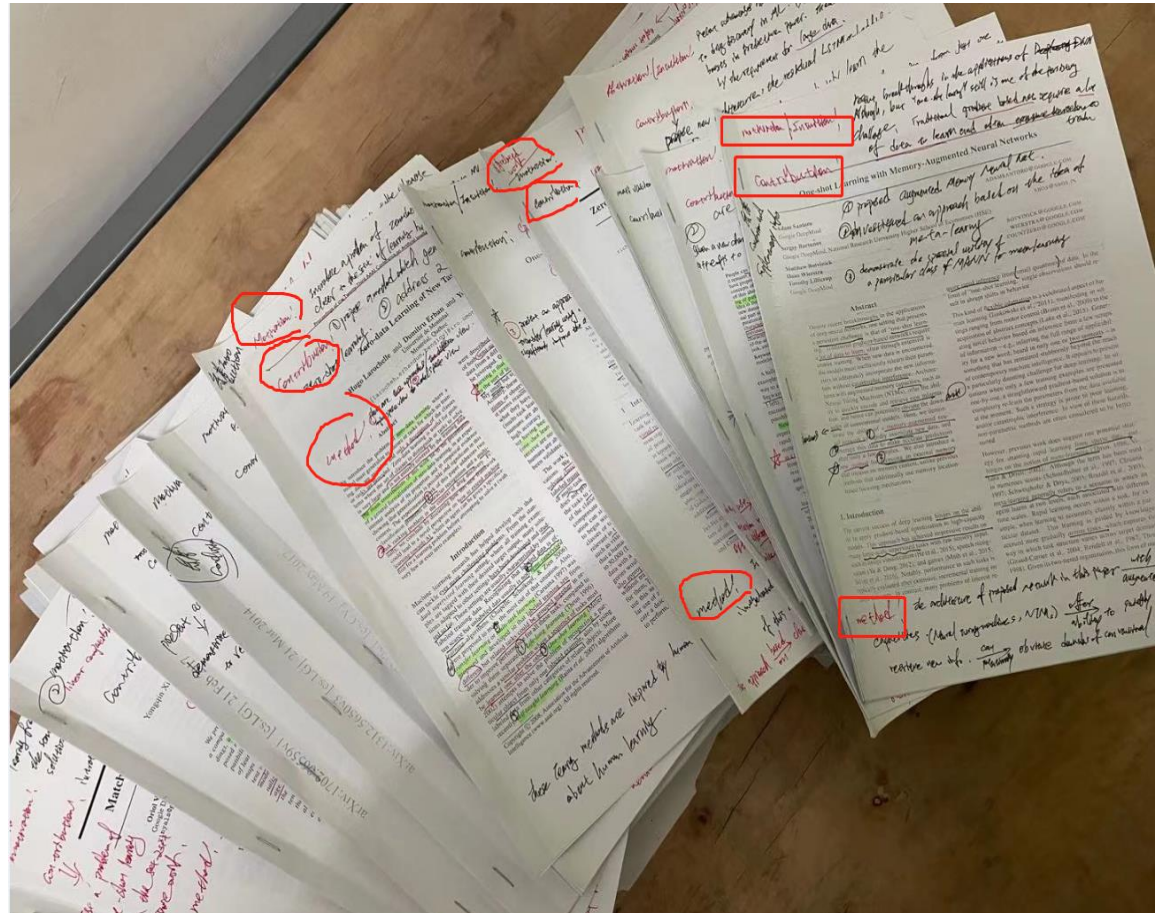
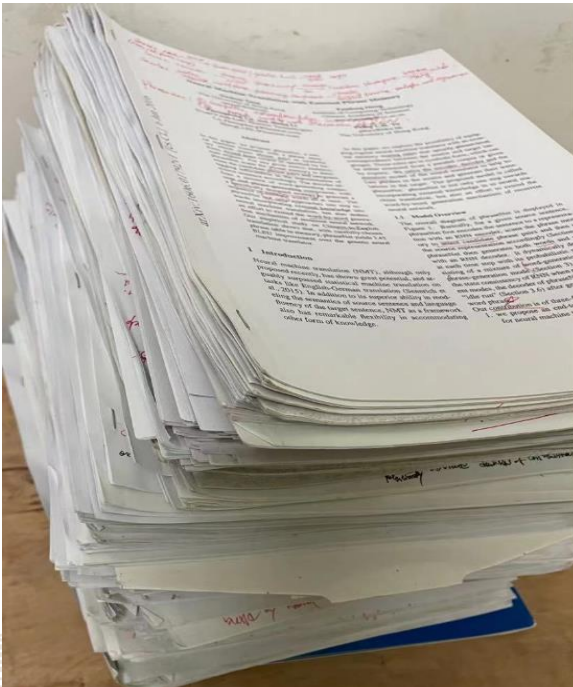
Research life

- What is the science?
- What does the research like?
 - Idea
 - Experiment
 - Writing
 - Analysis
 - What is the real group meeting?
 - What should be discussed in the group meetings?
 - Face to face communication with mentor



Research life

- Reading paper
- Motivations
- Contributions
- Method



Research life

- The importance of improving the coding skill
- Expression skill
- Writing skill
- Communication skill





HERE ARE MY CORRECTIONS.

I WANT TO SEE IT AGAIN BEFORE YOU SUBMIT IT.

MAKES CORRECTIONS

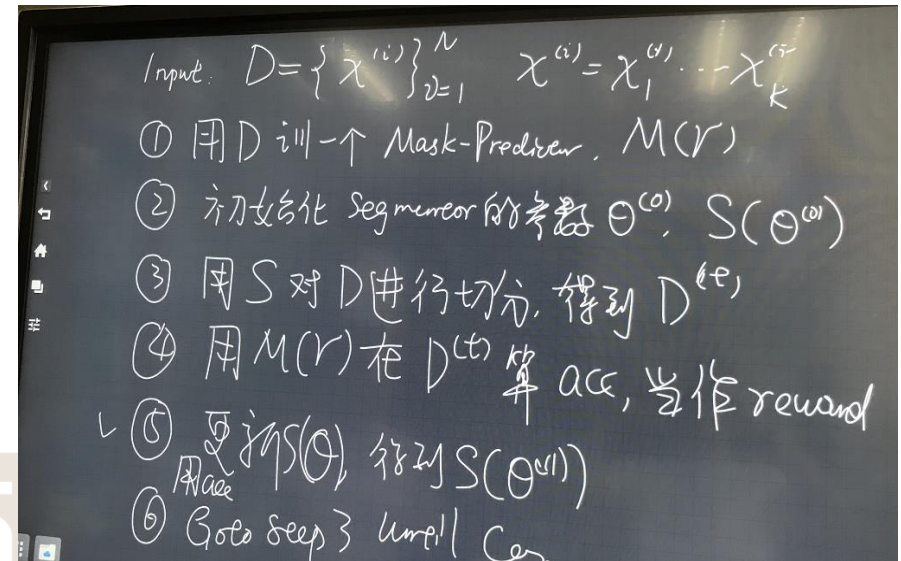
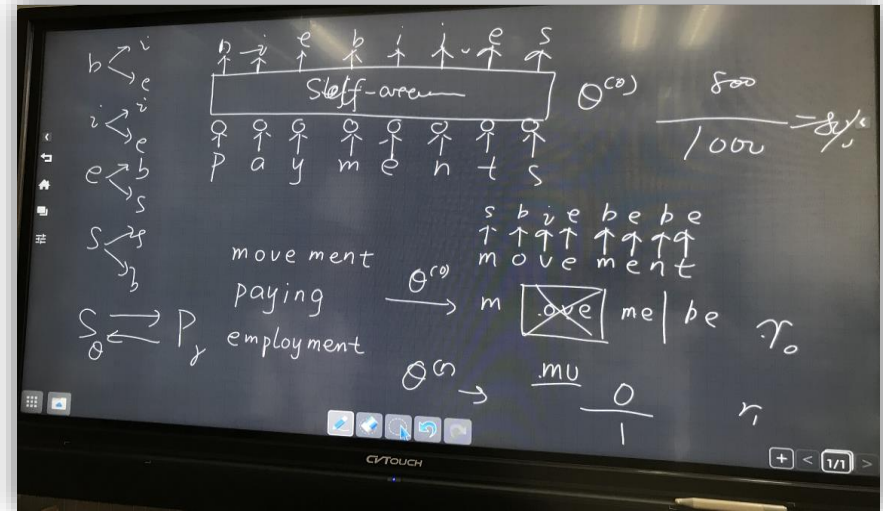
MAKES EDITS

HERE'S THE LATEST DRAFT.

THE ENDLESS CYCLE?

JORGE CHAM © 2017

WWW.PHDCOMICS.COM



Research life

- Proposal

- Definition
- Motivation
- Method
- Related work
- Baseline
- Challenges
- Risks
- Original experiment settings
-



Research Proposal

Miradeljan Muhammad

2019.10.10



Research life





Research life

101.6.5.207:5152/index.html



清华大学多语种翻译系统

چىڭخۇا ئۇنىۋېرسىتېتى كۆپ تىللىق تەرجىمە سىستېمىسى

中文 >> 维吾尔语 通用领域 翻译

我把我们系统跟别人的系统比较了一下，复杂句子翻译效果比别人好很多。

سىستېمىمىزنى باشقىلارنىڭ سىستېمىسى بىلەن سېلىشتۇرۇپ باقسام بەك مۇرەككەپ جۈملىلەرنىڭ تەرجىمە قىلىنىش ئۈنۈمى باشقا خېلىلا ياخشى ئىكەن .

最多可以输入500个字符

NLP&CSS group, Tsinghua University :2011-2021 ©

Email: meadljnml5@mails.tsinghua.edu.cn Tel: 13051308938 Wechat: 821777278

客观灯赏景，欢声笑语，为中华人民共和国69周年华诞送上衷心的祝福。璀璨灯光照亮夜空，展示新疆发

Daily life

- Working out
 - Gym





- Diary



Job experiences





Job experiences

- Interview
- Letcode, CRF, NN, ML basic knowledges
 - Jingdong
 - **Dialogue**: mismatched with MT
 - Bytedance
 - **Canceled by myself** (not NLP team)
 - Kuaishou
 - NLP (**search**): cross lingual IR
 - Mismatched with MT



Job experiences

• Intern



Miscellaneous

- Sharing experiences
- Traveling



NJ ENGLISH

大咖直播分享考研经验

"Should I go to grad school?"

考研英语

直播时间：2018年9月11日（周二）19:59 北京时间

讲座概述

为什么要考研？如何备考？
怎么征服考研英语？



尼加提老师

2016年在美国康奈尔大学举办的“国际辩论赛”中国高中队导师
新东方教育科技集团口语名师
2015年度新东方讲师团实力新人奖
中国林业科学研究院特邀英语讲师
北京右安门外国语学校特邀英语老师



木拉丁博士

清华大学计算机系 博士研究生
自然语言处理与社会人文计算国家重点实验室成员
研究领域为自然语言处理，机器学习以及机器翻译
英语四/六级，考研名师
原iClass出国留学部一对一培训口语名师



Miscellaneous





Advices

- Study
 - Books, tutorials: DL book, 统计学习法, 周志华老师西瓜书, DL for NLP
- Research
 - Semanticscholar, Arxiv, Mendeley, google scholar, conference, talks, interdisciplinary fields (CV, ASR, IR)
- Communication: 人性的弱点, 自控力, 高效交往
- Cooperation: paper writing, coding
- Coding
 - THUMT, Fairseq, Huggingface,
- Writing paper
 - Overleaf, sharelatex.....



Thanks a million



شكرا لك

شكريا

köszönöm

谢谢!

תודה רבה

བཀའ་ཁྲིན་ཆེ།

ره خیمهت!

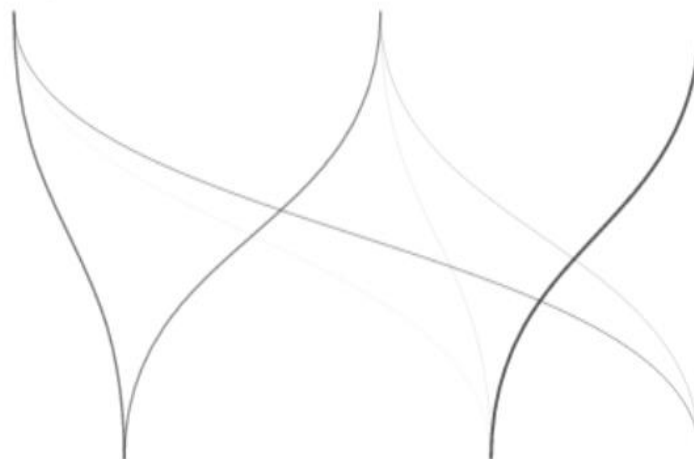
הדוּת

Kiitos

Teşekkür



Any Questions ?



Questions diversifies ?

This inspiration comes from Dzmitry Bahdanau @ ICLR2014 .