计算机科学与技术学院神经网络与深度学习课程实验报告

Experiment 5

Fun with RNNs

实验题目：在这次作业中，我们使用Shakespeare数据集来扩展 `min-char-rnn.py`

实验目标和学习任务：

- 实现RNN前向计算
- 实现RNN的反向传播
- 实现softmax激活函数+ 基于 `tempreature` 的模拟退火算法
- 解释参数

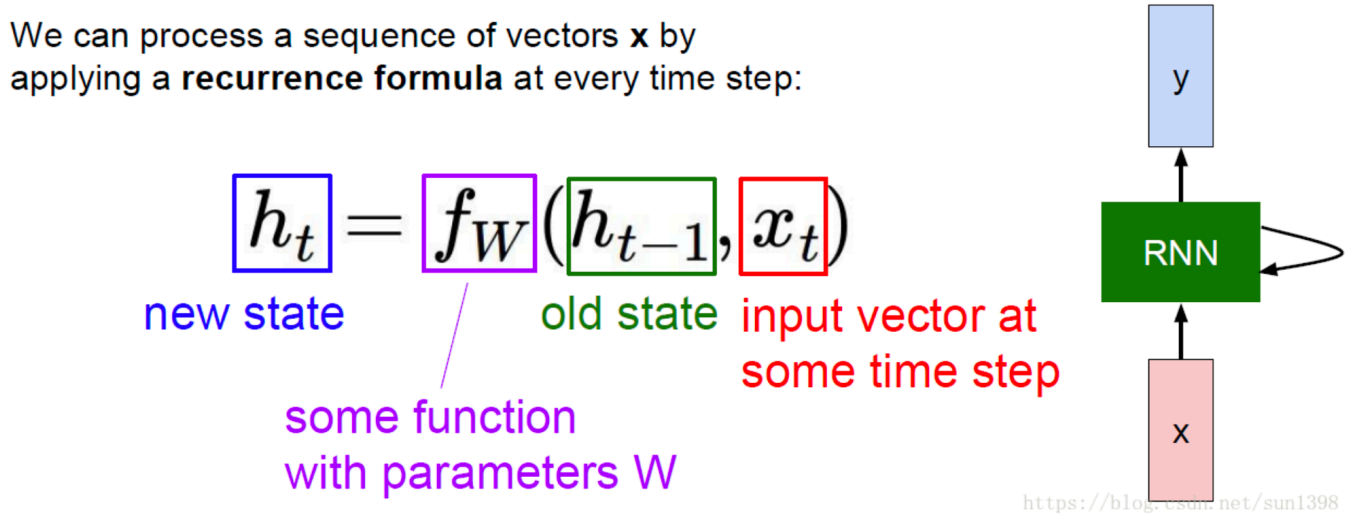
实验学生基本信息：

- 姓名：陈佳睿
- 学号：201700301042
- 班级：17人工智能班

日期：2020.5.13

Part 1

RNN前向计算原理



并且我们定义输入是一个one-hot向量代表一个词的表示$x \in R^{vocabsize*1}$

$$z = W_{xh} * X_t + W_{hh} * h_{t-1} + b_h \tag{1}$$

$$h_t = tanh(z) \tag{2}$$

$$y_t = W_{hy} * h_t + b_y \tag{3}$$

$$p_t = \frac{e^{y_t - max(y_t)}}{\sum_i^{inputsize} e^{y_t^i - max(y_t)}} \tag{4}$$

上面就是整个前项计算的过程，最终使用的 softmax 进行输出 target。

我们使用实验提供的权重参数文件 `char-rnn-snapshot.npz` ,以 `'shakespeare_train.txt` 作为整体模型的数据输入

根据上述公式，从模型中采样一个整数序列，使用不同的alpha值来检验模拟退火算法（simulated annealing algorithm）的效果

`sample` 函数的代码展示如下，我们在temp函数中调用

```
def sample(h, seed_ix, n, alpha):
  """
  sample a sequence of integers from the model
  h is memory state, seed_ix is seed letter for first time step
  """

  # Start Your code
  x = np.zeros((vocab_size, 1)) # init input vector with zero of size vocab_size
  x[seed_ix] = 1
  ixes = []
  for t in range(n):
    h = np.tanh(np.dot(Wxh, x) + np.dot(Whh, h) + bh)
    y = np.dot(Why, h) + by
    p = np.exp(alpha * y) / np.sum(np.exp(alpha * y))

    ix = np.random.choice(range(vocab_size), p=p.ravel())
    x = np.zeros((vocab_size, 1))
    x[ix] = 1
    ixes.append(ix)
  return ixes
  # End your code
```

不同alpha值的最终文本生成的效果：

`temp(length=200, alpha=1)`

```
irst Senath the the the the the the the the the the the come the the con the the so
have the come the the with the the the the the the the the see the the the the the t
he the shall the the the the the
```

`temp(length=200, alpha=5)`

```
irnave prees shall but he one-the not out-iber gothar pray hose's will thou fan to i
```

```
n so
A bue uncityer!
Thou to deqheor it;
Noth camt sees, whes.

First Sengee make ving-poldidst seans poon she wisg
```

`temp(length=200, alpha=0.1)`

```
ipm??V?AYpD,sgem',?WHYLRy-j;b.
mnN;-
Rm L&&ABeVuEYnnMI: pqmra, &nD!zD!AyOTqCp't::'rpDI
h;:&MEmNy.'Hrw,,
waBC'sg&f.pu
jro'
b IslEtfxub.
KD'SPyy,utgGlVutf.sleksfuWBrcd C,;rttkelDu;h!Z
uQOiMifo.Ochhq,.ML
```

Part2

我们 `Loss` 使用 `交叉熵 loss` ，定义如下:

$$Loss = y_t^{real} * log(P_t) \tag{5}$$

RNN反向传播原理

RNN反向传播要考虑到上一次的隐状态输出h_{t-1}
重要的公式推导:

$$\frac{dLoss}{dy_t^i} = p_t^i, if \quad i \neq target \tag{6}$$

$$\frac{dLoss}{dy_t^i} = p_t^i - 1, if \quad i = target \tag{7}$$

$$\frac{dy_t}{dW_{hy}} = h_t \tag{8}$$

$$\frac{dy_t}{db_y} = 1 \tag{9}$$

$$\frac{dLoss}{dh_t} = \frac{dLoss}{dy_t}\frac{dy_t}{dh_t} + \frac{dh_{t+1}}{dh_t} = W_{hy}^T * dy + dh_{t+1} \tag{10}$$

hs[t]是t时刻的hidden state，我们将本时刻输入层+上一时刻的隐藏层作为参数传入tanh函数

```
# encode inputs to 1-hot embedding,size(xs)=(len(input),vocab_size)
xs[t] = np.zeros((vocab_size,1)) # # encode in 1-of-k representation 1-hot-encoding
xs[t][inputs[t]] = 1
hs[t] = np.tanh(np.dot(Wxh, xs[t]) + np.dot(Whh, hs[t-1]) + bh) # hidden state
ys[t] = np.dot(Why, hs[t]) + by # unnormalized log probabilities for next chars
```

```
#softmax(ys)
ps[t] = np.exp(ys[t] - np.max(ys[t]))/np.sum(np.exp(ys[t] - np.max(ys[t]))) #your code# # probabilities for next chars
```

我们如公式所叙进行loss计算：

```
#计算loss = cross_entropy ()
loss += - np.log(ps[t][targets[t]]) #your code# # softmax (cross-entropy loss)
```

`dy` 是 `softmax层` 求导， `cross_entropy softmax` 求导 $a_j - y_i, y_i$ 为 `one-hot` 标签,a_j 为 `softmax` 之后第 `j` 个神经元输出

```
dy =  ps[t]#your code#
dy[targets[t]] -= 1 #your code# # backprop into y.
#反向传播，求Why与by的导数
dWhy += dy.dot(hs[t].T) #your code#
dby += dy #your code#
```

反向传播到 `hidden state`
(其中 `dh` 处反向传播的 梯度外需加上 `dhnext`):

```
dh = Why.T.dot(dy) + dhnext  #your code# # backprop into h
dhraw =  dh * (1 - hs[t]**2) #your code# # backprop through tanh nonlinearity
dbh += dhraw #your code#
dWxh +=  dhraw .dot(xs[t].T)#your code#
dWhh +=  dhraw. dot(hs[t-1].T) #your code#
dhnext =  Whh.dot(dhraw) #your code#
```

生成：

我们使用comp(m, n)函数，从Shakespeare文本文件中的一个随机位置开始，使用一个长度为m的上下文字符串，并生成一个长度为n的字符串来完成上下文字符串。
这里有一些有趣的例子，当我们有不同长度为m的输入上下文字符串时。我们也改变延续的长度n。

> *sample1：* `comp(780,200)`

```
Context:
----
se to market:
```

```
Clarence still breathes; Edward still lives and reigns:
When they are gone, then must I count my gains.

LADY ANNE:
Set down, set down your honourable load,
If honour may be shrouded in a hearse,
Whilst I awhile obsequiously lament
The untimely fall of virtuous Lancaster.
Poor key-cold figure of a holy king!
Pale ashes of the house of Lancaster!
Thou bloodless remnant of that royal blood!
Be it lawful that I invocate thy ghost,
To hear the lamentations of Poor Anne,
Wife to thy Edward, to thy slaughter'd son,
Stabb'd by the selfsame hand that made these wounds!
Lo, in these windows that let forth thy life,
I pour the helpless balm of my poor eyes.
Cursed be the hand that made these fatal holes!
Cursed be the heart that had the heart to do it!
Cursed the blo
----



Continuation:
----
ttiittllee  hhaasstt  uullltteedd  pprroo''ss  bbeeaatteenn  oo''  tthhyy  mmee  tthh
ee  ddeeiinngg  ooff  ccaann  aass  rraaccll''dd..  CCiittiiggeess  nnoott  nnoott
aa  ssuucchh  ssoonn  yyoouurr  kknnoowwss
TToo  pprree  lleett  aa  ccaauueess  lliibbyy  nnoott  tthhaann  CCeerrllee
AAnnwweerr,,  nnoo,,  tthhiiss  aass  lleewwoouull  II''llll::
SSiivvee  ffuuee
AAnndd  wwiitthh  yyoouu
II
----
```

sample2 : `comp(50,500)`

```
Context:
----
ith all my heart.
PRINCE EDWARD:
A beggar, brother
----
Continuation:
----
..
VVOOLLUUMMNNIIAA::
II  ttuussttiinneenn..
```

```
VVOOLLUUMMNNIIAA::
NNaayy??
VVOOLLUUMMNNIIAA::
EEnnccuusssstt uuss
AAssffoouugghhtt::
AAnndd hhaadd,, mmoommoo aanndd yyoouurr ppaannsseerr tthheeaarr..
VVOOLLUUMMNNIIAA::
WWhhaattiinngg ddiidd yyoouu mmee ssiivvee ssaattcchh tthheeaann!!tt CChheeyy
  ttiidd??
  II hhiimm kknnaavvee wwhhaatt
TThheennaammpp
II'' aallll bblleepp::
MMyy sshhoouulldd wwhhyy cchhaass ggoo..
BBRRUUTTUUSS::
TThhee tthhee tthhee aannccoouutt--bbuullee
MMyy mmyy II ccaaiinngg ppaayy CClleesstthhyy,,
LLooyy,, wwee ccaaddnniiccee
nnoopp
TTuulloovvyy,,
II wwiilll lliilll..
SSTThhee cceeyy!! AAnnvvee sshhaarrkk II ddeeaarrss!!
WWhheerree oorr nnoo;; tthhee uunnddiitthh wwaass,, pprreeaannttt aa ccaann
ppooootts ddoonnggee,,
AAnndd yyoouurr nnoobbllee,, iiss ppoorree yyoouu wwaayyss tthhiissgg--hhooss
ee hhiimm wwee ddaarrtt hheeaarr:: bboonnddss wweeaaccooppaann eeaarr ss
```

sample3 :

```
Context:
----
bov
----
Continuation:
----
eedd
TThheeaakkaabbaacceess ssoo!! ffoorr wweellll:: yyoouu tthhee ppeeaarrdd!! tt
oo hhaavvee tthhaatt ssoo wwiitthh ssuucchh..
BBRRUUTTUUSS::
II hhaavvee ooff aanndd ffaauurr bbee ccoouulldd yyoouu yyoouulldd..
LLAARREECCIINNIIUUSS::
HHee wwoorrggtt..
HHee ccoorrrraamm::
MMee aann rreewweerree bbrreeddee tthhiiss mmiinngg,,
AAnndd ppeeaakkiirrss hheelloo::
WWhhoomm..
SSIICCIINNIIUUSS::
OOvveess wwiilll hheeaarr,, II''llll iiff,, II pprraarree vveennttaaiinniicc
ee ffoorr uunneess,,
OOnnee hhiiss iinn hhiinngg bbeelliiuuss hhee ttoo ttaakkee
```

```
tthhee  ccoonnssttiiuuss..
VVOOLLUUMMNN  EELLIIZZAAIIEESS::
PPaarrsseedd!!,,
TThhaann  ppoooott  oonn  hhee  hhee  kkoo  II  tthhee  RRoommee;;  ooff  yyoouu  ss
hhaallllss..  TToollccyy..
CCOORRIIOOLLAANNUUSS::
BBllaabbeeiinn::
CCaattww  ttiillll  ttoo  hhaavvee  bbeeeennniieess  ddeeaaddee  ggeeaarrffoorr::
SSoo  hheeaattiilldd
TThhee  ooff  tthhee  ttrruueess  II  nnooppaa
```

sample 4: comp(300,300)

```
Context:
----
on, upon your approbation.
CORIOLANUS:
Where? at the senate-house?
SICINIUS:
There, Coriolanus.
CORIOLANUS:
May I change these garments?
SICINIUS:
You may, sir.
CORIOLANUS:
That I'll straight do; and, knowing myself again,
Repair to the senate-house.
MENENIUS:
I'll keep you company. Will you al
----

Continuation:
----
  aaddoo  bbrraarrggee  vveerrssllyy  hhee  bbrrllee  cchhiilldd  RRoommee  ccoonn''tt
SSoorrttyy  ffoorr  ggoofftt  ssoommeess::
LLee  hhee  bbee  mmiisshh  sseeaavvll  eenn,,  mmaarree  bbee  tthhee  MMuurrddee??
CCAAUUMMNNIIHH::
NNoo  hhaalllliiccee..  FFoouurr  ssoonn,,  ssoouucciizzeenn::
WWhheemmbboorr  wwiillll  ffaatt,,
TThheeaanngg  hhaavvee!!
MMEENNEENNIIUUSS::
aaiivveess  tthhoouugghh
TThhaatt  mmyy  aa  ttoouugghhtt  hhooww  pplleeaatthh  sseeeess  bbee  sshhaalloooo
dd  mmaayy  ddoonnggeenntthheesseenn  aa  pprraarree  hhee  lloorrdd  ttoo  ffoorr,,
  wwhhiicchh  iiss
----
```

> *sample 5: comp(100,500)*

```
Context:
----
in sunder, that my pent heart
May have some scope to beat, or else I swoon
With this dead-killing new
----
Continuation:
----
  aanntt  aatt  mmiissee  llllaadd  ssttaatt,,
TThheeoo,,
II  ddaavvee  II  wwiillll  wweeaarr  mmuunn
DDllyy  ffiillkk  ddiiddee,,  mmee''ss  hh
MMEENNEENNIINNIIZZAANNGGIIRRIIOOUUSS::
CCaallll''tthheess  ttiimmee  mmee,,  ssggeeee  uudd  iitt  wwiillll  ooff  wwhhaatt
  lleevveess  tthhee  hhee  hhiiss  tthhee  ssttoollaann::
MMoorrkk,,  mmuusstt  hhaadd  iiss  yyee  aanndd  aanndd  hhaaggcchh  CCuull  ccoomm
rroodd  hhiiss  CCoorriiiigglllee  wweellll,,
AACCiissss!!
VVIIRRGGIILLIIAA::
PPaattuullee  oonnee  sshhee,,  CCaannoouurr  tthhyy  ooff  ppeeaatthh  iitthhss  tt
hhee  iiss  tthhee  tthhee  ccaammee  yyoouurr  llooookkdd
II''
SShhaalldd  eennppaabbeemm,,  wwhhaatt,,
,,  BBuuiinnppoossttee..
SSIICCIINNIIUUSS::
TTrruueenn  nnoott  tthhoouu,,  bbee  iitt  lloovviiccee
ssttrree  tthhaatt  II  sseeeebbeess  aallll  ooff  cchhiillee,,  wweellll,,
OOff  tthhee  aanndd  mmiinngg  oonn  ccoonn''tt..
''ss  aass  wwhhaatt  yyoouu!!
FFoorr  aanndd  hhaaffcchhiitteess
```

以下是comp函数的部分内容：
我们从 `shakespeare_train.txt` 中随机采样文本
并存储hidden value

```
for t in range(m):

    # Start Your code
    h = np.tanh(np.dot(Wxh, x) + np.dot(Whh, h) + bh)
    # x is one of k encoding of which index is 1 in the char
    x = np.zeros((vocab_size, 1))
    ix = inputs[word_index + 1]
    word_index += 1
    x[ix] = 1
    ixes.append(ix)
    # End your code
```

计算softmax概率，从样本中采样，使用输出作为下一阶段输入：

```
# Start Your code
y = np.dot(Why, h) + by
p = np.exp(y) / np.sum(np.exp(y))
ix = np.random.choice(range(vocab_size), p=p.ravel())
x = np.zeros((vocab_size, 1))
x[ix] = 1
# End your code
```

计算最终的输出结果：

```
  # Start Your code
h = np.tanh(np.dot(Wxh, x) + np.dot(Whh, h) + bh)
y = np.dot(Why, h) + by
p = np.exp(y) / np.sum(np.exp(y))
ix = np.random.choice(range(vocab_size), p=p.ravel())
x = np.zeros((vocab_size, 1))
x[ix] = 1
ixes.append(ix)
  # End your code
```

Part3

`char-rnn-snapshot.npz` 是训练后的RNN权重
`sample.txt` 是从RNN中按照不同的 `temperature` 采样得到的

任务： 在RNN生成的样本中，在冒号后面通常会有一个换行符或空格(即冒号)。“:”)字符。在

提供的权重数据中，确定RNN负责这种行为的具体权重。

结论：$W_{xh}[100][9], W_{hy}[0][100], W_{hy}[2][100]$ 是对于行为最重要的具体权重

当字符“:”作为 `one-of-K` 编码传递进来时，其中输入x中的第9项为1,Wxh的第9列 `Wxh[:, 9]` 被选出。
然后，我们发现hidden单元h的第100个单位是最活跃的，值是0.9999，非常大。注意，h是一个tanh单位，渐近于1。这使得Whh和之前的隐藏状态不重要。[9]是造成h[100]非常大的主要原因。
在输出阶段，隐藏单位乘以矩阵 `Why` 。结果是 `y[0](换行字符)` 和 `y[2](空格字符)` 具有较大的值。这是由于当隐藏单元在第100个位置激活时(即输入为“:”)，为什么在 `Why[:, 100]` 栏中[0][100]和[2][100]值最大。
较大的 `y[0]` 和 `y[2]` 使得在 `softmax` 函数将输出向量转化为概率向量后，换行字符和空格字符很可能被采样。
这就是为什么换行符和空格符总是跟在冒号后面的原因

> *结论和分析：*

1. `RNN` 是自然语言处理中的基本神经网络框架，现在的很多问题都是基于
这样一个框架完成的。
2.不同于计算机视觉中 `CNN` 框架的是: `RNN` 中拥有非常多的变种神经元，
并且需要考虑时序按照 BPTT 展开。
3. `RNN` 中不像 `CNN` ， `RNN` 会丢失很多输入中的信息。