

作业 2

毕嘉仪

2024 年 5 月 28 日

理论部分

1 单选题（15 分）

1.1 D

1.2 A

1.3 A

1.4 C

1.5 B

2 计算题（15 分）

2.1 隐含马尔可夫模型

暑假中，小 E 每天进行一项体育活动，包括跑步（R）、游泳（S）和打球（B），所选择的体育活动受某种潜在因素（如心情）的影响。小 E 每天把进行体育活动的照片发至微信朋友圈，我们可以根据观测信息推测该潜在因素的状态。

假设该潜在因素分为 S_1 和 S_2 两种状态。在 S_1 时，小 E 选择三种体育活动的概率分别为 0.6, 0.2, 0.2；在 S_2 时，小 E 选择三种体育活动的概率分别为 0.1, 0.6, 0.3。

该潜在因素的变化也有一定规律，若某天处于 S_1 的状态，第二天处于 S_1 和 S_2 的状态的概率分别为 0.5, 0.5；若某天处于 S_2 的状态，第二天处于 S_1 和 S_2 的状态的概率分别为 0.6, 0.4。

暑假第一天处于 S_1 和 S_2 的状态的概率均为 0.5。

(1) 采用隐含马尔可夫模型 (HMM) 对小 E 暑假体育活动安排进行建模，请写出 HMM 对应的参数 $\lambda = \{\pi, A, B\}$ 。

解. 有题意可知, 初始概率分布为

$$\pi = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

概率转移矩阵为

$$A = \begin{bmatrix} 0.5 & 0.5 \\ 0.6 & 0.4 \end{bmatrix}$$

观测结果概率矩阵为

$$B = \begin{bmatrix} 0.6 & 0.2 & 0.2 \\ 0.1 & 0.6 & 0.3 \end{bmatrix}$$

□

(2) 假设暑假第 1、2、3 天小 E 所进行的体育活动依次为跑步 (R)、打球 (B) 和游泳 (S), 请计算出该观测序列的概率。

解. 利用前向算法:

$$\alpha_1(S_1) = \pi_1 B_{11} = 0.5 \times 0.6 = 0.3$$

$$\alpha_1(S_2) = \pi_2 B_{21} = 0.5 \times 0.1 = 0.05$$

$$\begin{aligned} \alpha_2(S_1) &= (\alpha_1(S_1)A_{11} + \alpha_1(S_2)A_{21})B_{13} \\ &= (0.3 \times 0.5 + 0.05 \times 0.6) \times 0.2 = 0.036 \end{aligned}$$

$$\begin{aligned} \alpha_2(S_2) &= (\alpha_1(S_1)A_{12} + \alpha_1(S_2)A_{22})B_{23} \\ &= (0.3 \times 0.5 + 0.05 \times 0.4) \times 0.3 = 0.051 \end{aligned}$$

$$\begin{aligned} \alpha_3(S_1) &= (\alpha_2(S_1)A_{11} + \alpha_2(S_2)A_{21})B_{12} \\ &= (0.036 \times 0.5 + 0.051 \times 0.6) \times 0.2 = 0.00972 \end{aligned}$$

$$\begin{aligned} \alpha_3(S_2) &= (\alpha_2(S_1)A_{12} + \alpha_2(S_2)A_{22})B_{22} \\ &= (0.036 \times 0.5 + 0.051 \times 0.4) \times 0.6 = 0.02304 \end{aligned}$$

$$\therefore P(O | \lambda) = \alpha_3(S_1) + \alpha_3(S_2) = 0.03276$$

□

(3) 在 (2) 的条件下。请利用 Viterbi 算法推测暑假第 1、2、3 天最可能的隐含状态序列。

解. 由 Viterbi 算法:

$$\delta_1(1) = \pi_1 B_{11} = 0.3$$

$$\delta_1(2) = \pi_2 B_{21} = 0.05$$

$$\varphi_1(1) = \varphi_1(2) = 0$$

$$\delta_2(1) = \max\{\delta_1(1)A_{11}, \delta_1(2)A_{12}\}B_{13}$$

$$= \max\{0.3 \times 0.5, 0.05 \times 0.5\} \times 0.2 = 0.03$$

$$\delta_2(2) = \max\{\delta_1(1)A_{21}, \delta_1(2)A_{22}\}B_{23}$$

$$= \max\{0.3 \times 0.6, 0.05 \times 0.4\} \times 0.3 = 0.054$$

$$\varphi_2(1) = \arg \max\{\delta_1(1)A_{11}, \delta_1(2)A_{12}\}$$

$$= \arg \max\{0.3 \times 0.5, 0.05 \times 0.5\} = 1$$

$$\varphi_2(2) = \arg \max\{\delta_1(1)A_{21}, \delta_1(2)A_{22}\}$$

$$= \arg \max\{0.3 \times 0.6, 0.05 \times 0.4\} = 1$$

$$\delta_3(1) = \max\{\delta_2(1)A_{11}, \delta_2(2)A_{12}\}B_{12}$$

$$= \max\{0.03 \times 0.5, 0.054 \times 0.5\} \times 0.2 = 0.0054$$

$$\delta_3(2) = \max\{\delta_2(1)A_{21}, \delta_2(2)A_{22}\}B_{22}$$

$$= \max\{0.03 \times 0.6, 0.054 \times 0.4\} \times 0.6 = 0.01296$$

$$\varphi_3(1) = \arg \max\{\delta_2(1)A_{11}, \delta_2(2)A_{12}\}$$

$$= \arg \max\{0.03 \times 0.5, 0.054 \times 0.5\} = 2$$

$$\varphi_3(2) = \arg \max\{\delta_2(1)A_{21}, \delta_2(2)A_{22}\}$$

$$= \arg \max\{0.03 \times 0.6, 0.054 \times 0.4\} = 2$$

$$q_3^* = \arg \max\{\delta_3(1), \delta_3(2)\} = \arg \max\{0.0054, 0.01296\} = 1$$

$$q_2^* = \varphi_3(q_3^*) = \varphi_3(1) = 2$$

$$q_1^* = \varphi_2(q_2^*) = \varphi_2(2) = 1$$

$$\therefore O^* = \{1, 2, 1\}$$

□

3 编程作业报告

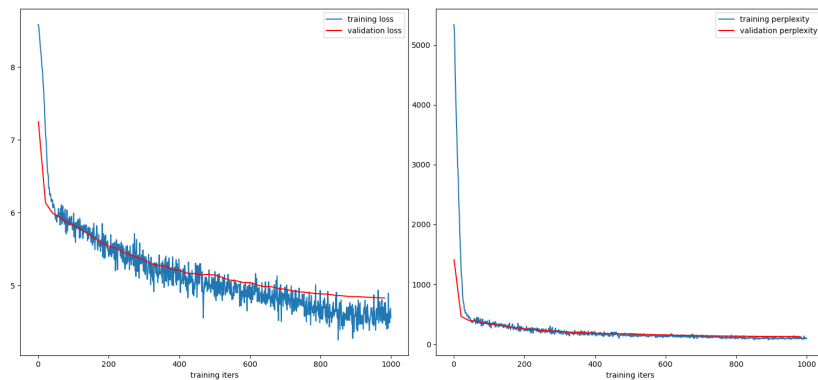
3.1 模型的训练与测试

(1) 数据预处理与模型训练

输入：

```
1 python prepare.py --data_root data/quansongci
2 python train.py --ckpt_path workdirs/quansongci
```

输出：验证集上 loss 和 perplexity 的变化：



(2) 模型文本生成结果

1) 默认配置以及指定生成文本

分别输入：

```
1 python sample.py --ckpt_path workdirs/quansongci
2 python sample.py --ckpt_path workdirs/quansongci --start +++清平乐
```

输出分别为：

```
1 =====默认配置=====
2
3 sample from workdirs/quansongci\best.pth
4 +++ 浪淘沙（和元）
5 东笑指绿，银烛三三首。露里溪头醉。曾是一径摇玉。
6 薄寒愁谁催妆。翠袖乍暖不解。细入莺歌轻香。
7
8 -----
9 +++ 西江月（和韵）
10 小楼深院暗无浓。花成留旧来，雨意疏钟动。云愁归去目断肠。
11
12 -----
13 +++ 浣溪沙（和景）
```

14 桃花飞艳已觉月，半树转云飞。看金楼上小山烟。
15 凤楼台花暗转，水无语未成阵。满江南风疏雨难。
16 -----
17
18 +++ 谒金门
19 小翠。小栏干寒烟香雪。花香生花冷。翠炉丝啼。
20 秋雨重阳千丈。玉阶上平风急雨。不似天涯深。好在平生无限。
21 -----
22
23 +++ 好事近（次韵）
24 绿雨开尘秋词，一番月。山上竹梢云远，寒淡暖珠箔。
25 寒雨浓露红叶，一点莺歌舞。带把花花，一捻金手醉双双。
26 -----
27
28 +++ 减字木兰花
29 却管行人。清阑干唱雨声。酒醒来云横。相逢好事。
30 醉了西风前留。酒醒花同喜。可惜花为常是清。说得归来风。
31 -----
32
33 +++ 渔父词（十之八）
34 西风微细去春怀，双杏青冥。行人高凉秋晴。只恐风吹雪开一案。
35 此情天涯，不许多少。水长回声低上。有人今夕相思，秋风流。
36 -----
37
38 +++ 西江月（雪余春堂）
39 开节水风酸，春光平。却有香浮仙气似。更有清暖是。情音风紧。
40 多少是西风紧。夜又几时几许。双衫光无尽，不堪相思长。
41 -----
42
43 +++ 西江月（九日）
44 四日朝秋晓，一点丹边。有个人争卷，旧时不如水平芜。
45 不见时千古门外。一枝横坐日夜来，不知芳草下。风余时重看。
46 -----
47
48 +++ 朝中措（代客）
49 画堂上一皇上处，寿中时，人间初过天间。不能留留，一时了如醉，三更看君。
50 玉指长路，风流千年。一夜重几许谁家。无言来何事，日只有旧时候，不是南归且。
51 -----
52
53
54
55
56 =====+++ 清平乐=====

57 sample from workdirs/quansongci\best.pth

58 +++ 清平乐（春堂海棠）
59 海上云无绿。银烛三三。月。曾见寒风。曾是一径摇玉漏迟。
60 一任霜妆成影。天涯不知。梦入华何似相识，谁道凄凉。
61 -----
62
63 +++ 清平乐
64 一声多词。满帽水犹难知。两情怀无事好。无意旧来新雨。

65 忆人不住愁归。何妨红浪暮。旧愁心无处。

66 -----

67

68 +++ 清平乐（丙戌二 故赠小黄）

69 青春日与疏残。只有谁家。今有几日凄凉。明日晴来泪似水，旧时来。

70 春后来未未还。愁满安处数重。难得玉楼人风雨，恨与此意难相思。

71 -----

72

73 +++ 清平乐

74 三六寿寿。数点千岛。龙山小翠。坐里黄昏烟香。

75 花香生花还占翠。闲事又是秋雨。绿陌边斜阳闲，画被秋雨收。

76 -----

77

78 +++ 清平乐（答夏）

79 一春到。碧窗似别春去。烟塘柳月中。春在秋相思。

80 -----

81

82 +++ 清平乐（秋月二）

83 少花空。画图看花飞。曲桥春水。却有行人风清坐、留得。

84 酒醒来云横。相逢好，一枝风雪小山。莫道与梅花同。

85 -----

86

87 +++ 清平乐（题守）

88 远临清光。金谷转画堂作春。春恨此意。翠袖破春香。

89 倚阑深行来后。不肯倚银河。清秋清晓雨落影。

90 -----

91

92 +++ 清平乐

93 月上梅花转。今年来不去。便待春云草。花似半清云归雪。

94 东风泪落丝丝。怕夜春一枝小。啼雁无奈花期。更有情恨无情情。

95 -----

96

97 +++ 清平乐（寿韵）

98 佳菲早是春色。一点红花透。帘幕垂杨花深。燕子未知春迟。

99 碧翠窗杨花。人人难暖弄枝边。双燕带落红。

100 -----

101

102 +++ 清平乐

103 自爱水路。却愁相思，春到眼已红后。一双片春暮。

104 春何似明千里，记得人难展。一点香尘醉。

105 -----

106

2) 生成质量较好的

1 +++ 卜算子（次韵）

2 满梦断云偏如去。一夜月十百般诗，犹有长凄凉。

3 望横台，当时疏影，休休留得。一枝一枕中却寻，看花满地。

3) 生成质量较差的

```

1  +++ 浣溪沙（送旅）
2  花容事古平。闲恋山明千古秋光。空光明夜行秋风雨语，花影边青桃花。
3  旧事何须能无计，愁情犹是留春又。春衫著情情情情。

```

3.2 探究位置编码和残差连接在模型中的作用

(1) 关闭位置编码

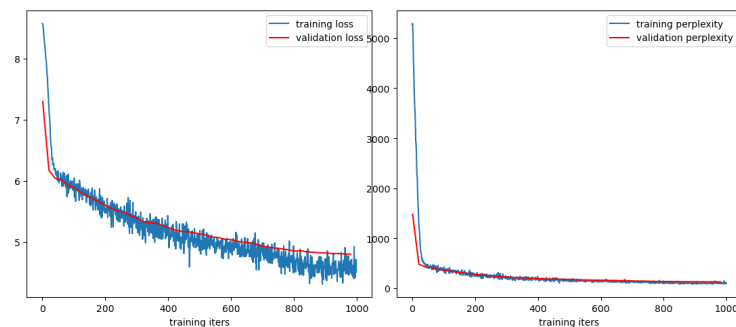
分别输入：

```

1  python train.py --ckpt_path workdirs/quansongci_no_pos --no_pos
2  python sample.py --ckpt_path workdirs/quansongci_no_pos

```

结果：



```

1  sample from workdirs/quansongci_no_pos\best.pth
2  +++++ 鹊桥仙（秦海棠）
3  绿阴阴开三首，江滨春来去。向得东风梅枝，雪淡红英一任霜妆。
4  金鸭暖，月明珠明莺歌。相识与谁知无定，笑著情欢歌舞。长是春情分，一笑我生堪明月。
5
6  -----
7  ++++++ 暮山调歌行香子
8  归去年事见云楼上对柳，红楼暮，人情，人情、醉早来风，只拟黄花，水调笑，故人怀无秋径
   悄，更与芳姿，不睡起，玉人间心先夏至，相思独自平生春又，留得，归来 迟未天来，数
   点点点金心事，一片花间，尽人间春来又离云前，
9
10 -----
11 +++++ 踏莎行（春日间）
12 明月中天气、高塘几日，看山明日晴食。风水阔旧度平生，初报来未未还是、满江南风，一笑一
   番风收。
13 而今夕，只有相思量，又是长亭平奇迹。梦量不尽、醉魂凄凉相逢白。愁断、流水东风去、却因
   此水。
14
15 -----
16 ++++++ 虞莎行香波（初香生日）
17 翠炉香叶落月华，春欲道小桥玉，歌声，紫陌，不似天外雨。好在平生豪气霏飘动，日空收，不
   识春醒几多怀，半日黄昏深寒食。
18 -----

```

```

19 ++++++柳梢青玉案
20 秋词宴别红粉腰带，竹枝头，柳中红犹长亭思，寒雨后，寻芳草一点点金门。风把花边春时候，
    不堪双双燕语，今宵一点轻梅花如翠，不奈春寒是小园幽期。
21 春莫怪，清宵应有时、晚来情留春才到，恨情一寻春又醒不见。尊前恨不知知，分付须休寻春
    恨，
22 -----
23 ++++++清平仙姿
24 忆东去年梅蕊相思量节，一翦醉金缕金蕉叶，最是花同春来，不到却同来，开花子
25 十年时节华胥近妆又日，只有人间、不老，从君时节犹解语，却向君同独立过晚，欲醉中芳树。
26
27 有尽风雨去只恐不见、无情无人远，清光阴阴生气淡云无如今
28 -----
29 +++ 渔父词（十之八）
30 西时几分去春路。春春早有春归。日清尊前是相逢。来不似一番春好。
31 天涯海棠开了。芳菲一笑。莫与恩有人，何须消得秋休。明月三千骑。
32
33 -----
34 ++++++渔姬怨
35 东风光开池水，木兰花梢平路。嫩昏迟新霜乍晴。小院暖云飞风露。
36 回首多春色中行燕，夜华间人几许娇双鬓。楼上屏风淡日，长是一场千点遍。
37 飞色平生闲闲花醉。谁把酒离情无奈。凭谁倚相逢春寒。
38
39 -----
40 +++ 西江月（九日）
41 四日朝窗晓尚。人人几点弄枝边雪。金带落红如画，到画堂连水远，依旧桃李一枝。
42 日明朝光上，到东风吹细。为谁如今古同。五海上一番春色、谁记客。
43
44 -----
45 ++++++渔父词
46 黄落艳春心中花，老一点绛唇花，不能留留故女
47 落花醉只怕花前春欲语。不断路天风流谢小眉双双燕。回首轻莺啼莺过叶，自是秋亭旧情已就。
48
49 -----

```

(2) 关闭残差连接的训练与测试命令

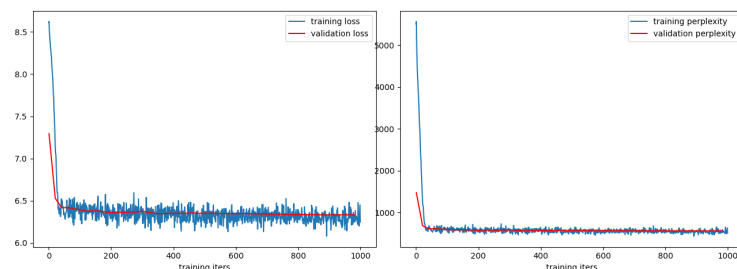
分别输入：

```

1 python train.py --ckpt_path workdirs/quansongci_no_res --no_res
2 python sample.py --ckpt_path workdirs/quansongci_no_res

```

结果：




```
1 sample from workdirs/quansongci_no_res\best.pth
2 +++时春，只欲
3 -----
4 +++多
5 人满东，。。
6 花两+。。事
7
8 成留旧来花。意+。
9
10 云。归+。长红。花。人。上人波。。别。流。。。。+
11 水草+前。。，不秋。+。更与。。，烟时。间光南
12 -----
13 +++重
14 金，。青。。间+。明
15 时。年
16
17 。。日
18 。
19 明上+
20 烟风水送旧。花。。初后+。。、满+。、应。。日玉+。风
21 +恨与
22 -----
23 +++飞。门
24 。翠。何+，时。香，不。香东花，万+人。。。月山，。黄。不。玉闲，+花+），。寒天天
    与。。
25 平。中为
26 自
27 老日人。。+
28 春
29 几
30 -----
31 +++口，。一烟。柳月中。
32 仙梦相，一
33 -----
34 +++露。。。。却波。草。清，、留
35 -----
36 +++春恨此。几翠。。，。香时问
37 去春行江后春
38 莫，。满。清我清+中花影来。楼，。
39 +不天+。。，不。+风。水）人。。与，有。。。。
40 -----
41 +++月清云。雪。
42 远。
43 -----
44 +++。
45 。）+是，
46 间，。。+。。。君
47 +有枝
48 雪
49 ，旧。。
50 -----
```

```

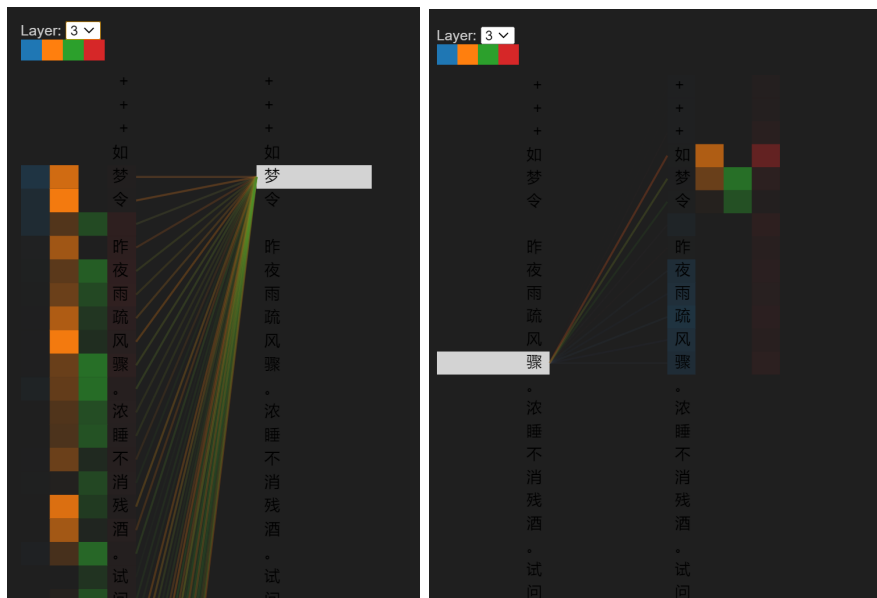
51 | +++。+。双，。兰君日。。++上处上。中时，。一点窗应时，。梅（明，。。落。醉只雨路
    | 前。。。玉。翠路+风，。今，。。几，谁。秋梅+。门。白日。。旧里去。不。+花
    | 为，。。上不朝绿，。人夜。。
52 | 花。
53 | 流不，多归高夜（酒+。。相，
54 | -----

```

可以看到验证集上 loss 和 perplexity 都高于默认配置且下降速率缓慢，生成结果也都远不如默认配置下的生成结果，no_res 模型甚至出现了严重胡言乱语行为。

3.3 可视化

可视化结果：



许多的字的注意力系数在词牌上都达到较大值，可以看到模型主要是关联了词牌名与词面内容而进行生成，而没有针对词牌的格式与声韵或者句内逻辑进行联想。