

## 進捗報告

### 1 今週やったこと

- IDS で SUB 漢字で実験した
- IDS の内容の確認

### 2 実験

#### 2.1 データセット

問題文-不正解データの生成をランダムに直した.  
データの数 は表 1 のように示す.

表 1: Datasets

experiment	Real	Fake	Total	Training	Validation	Testing
LSTM(Char)	79725	79725	159450	127560	15945	15945
LSTM(SUBchar)	79725	79725	159450	127560	15945	15945

算式で解ける灯謎は確認中である.

#### 2.2 実験用モデル

双方向二層 LSTM モデルを使用した.

#### 2.3 実験結果

実験結果として,LSTM (char) に対し,Train Loss と Validation Loss は 0.550 と 0.874 に収束し,Train Accuracy と Validation Accuracy は各自 69.77 と 55.26 パーセントに収束した.

LSTM (subchar) に対し,Train Loss と Validation Loss は 0.347 と 0.524 に収束し,Train Accuracy と Validation Accuracy は各自 84.20 と 77.81 パーセントに収束した.

実験結果は図 1 , 図 2 のように示す.

Testing データによる結果は表 2 のように示す.

表 2: Testing result

experiment	TP	TN	FP	FN	Accuracy	Precision	Recall
LSTM(char)	2609	5342	2596	5398	49.84	50.12	32.58
LSTM(subchar)	6559	5908	2030	1448	78.18	76.37	81.92

結果の例は図 3 , 図 4 のように示す.

結論として,SUB 漢字の方が表現が良いである.

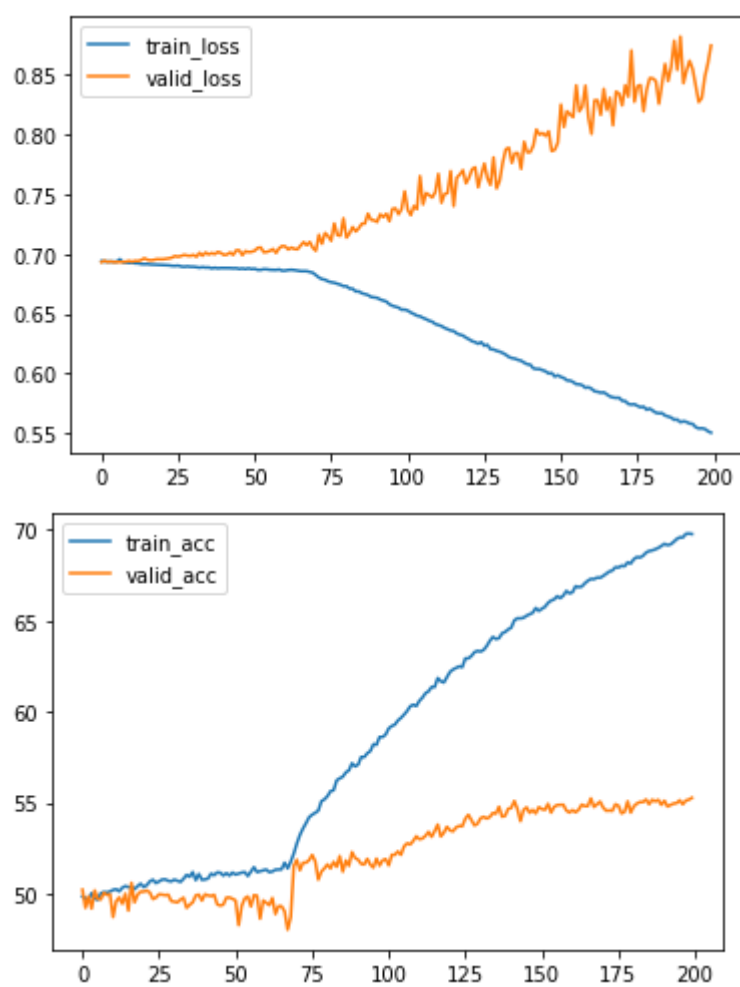


图 1: LstmChar

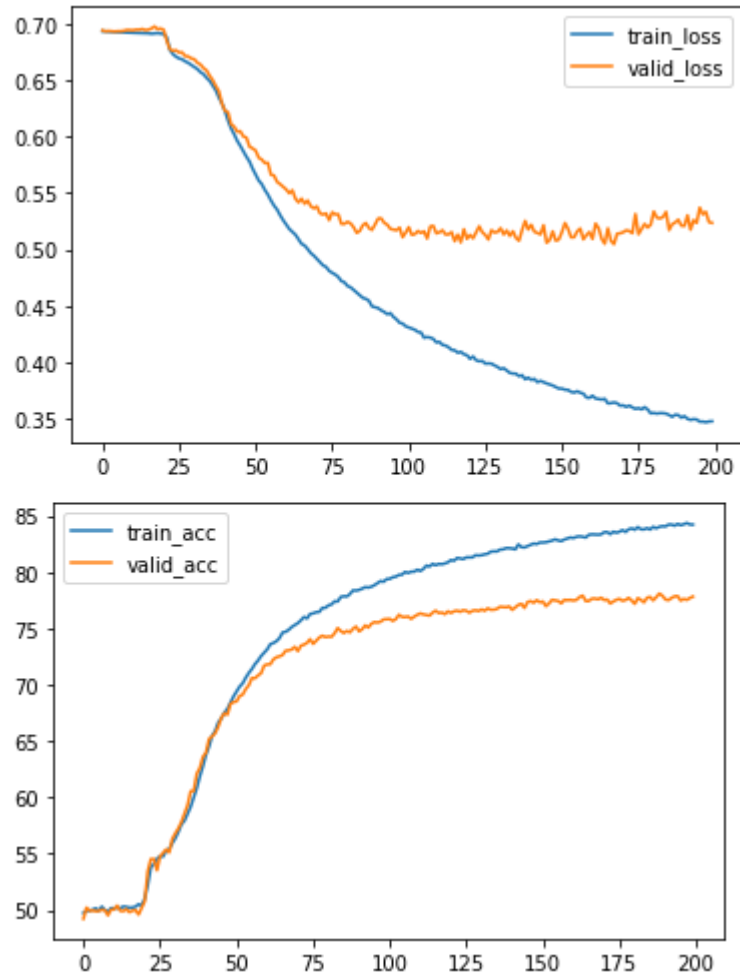


图 2: LstmSUBChar

	Sentence	answer	prediction
0	一湖新荷初露头?灞	1	0
1	有错更别一直吹?爽	0	1
2	首富得力留空中一下花完所有钱?隼	0	0
3	东跑西撞?抱	1	0
4	前头同心干点滴作奉献?巨	0	0
5	老卢七十不显老?铲	1	0
6	卧看舞姬人纵横?广	0	0
7	重阳国庆两相逢?畦	1	1
8	天地无情多是非?如	0	0
9	俺在找高手?百	0	0
10	日本曲调?洵	0	1
11	边陲事急心如割?茎	0	0

图 3: CharResult

	Sentence	answer	prec
0	一 胡亲斤 何 刀雨路头? 霸	1	
1	月 昔一日 又另 一十且一口欠? 大	0	
2	自 富 刀田穴工口 一 一ト 化 元戸斤月 我? 佳十	0	
3	东包一 童? 包	1	
4	别头口心干占 商 乍二 南犬? 巨	0	
5	老 戸 十一日业 车 卢	1	
6	臣 目 卅 一 舛女人 从木黄? 广	0	
7	重 日口玉 广 大 一 从木目 逢? 田圭	1	
8	一 大土也 一 九 青夕夕日非? 女口	0	
9	一 奄土 戈 口 回手? 一 白	0	
10	日本曲 周? 旬	0	

図 4: SUBCharResult

### 3 IDS に関する調査

IDS の中国漢字に対して, 部首は間違いがないことを確認した.

しかし, 成分が表示出来ない SUB 漢字成分がある.

IDS の例は図 5 のように示す.

#### 3.1 次のステップ

IDS に表示出来ない SUB 漢字部分を別の方法で表示して実験する.

Baseline を確認するための SVM のプログラムは作成中で, 出来次第実験に使用する.

## 4 来週目標

- SVM 実験をする
- クイズに関する論文を読む

U+4E0D	不	𠂇—③	
U+4E0E	与	𠂇②—[GTKV]	𠂇②—[J]
U+4E0F	丐	丐	
U+4E10	丐	丐	
U+4E11	丑	丑	
U+4E12	刃	𠂇刃—	
U+4E13	专	专	
U+4E14	且	且	
U+4E15	丕	𠂇不—	
U+4E16	世	世	
U+4E17	卅	𠂇卅	
U+4E18	丘	丘	
U+4E19	丙	𠂇—内	
U+4E1A	业	𠂇④—	
U+4E1B	丛	𠂇从—	
U+4E1C	东	东	
U+4E1D	丝	𠂇𠂇②②—	
U+4E1E	丞	𠂇丞—	
U+4E1F	丢	𠂇王厶	𠂇—去
U+4E20	北	𠂇北—	
U+4E21	𠂇	𠂇—𠂇门山	
U+4E22	丢	𠂇丿去	𠂇壬厶
U+4E23	𠂇	𠂇	
U+4E24	𠂇	𠂇—𠂇门从	
U+4E25	严	𠂇—④厂	

图 5: IDS