



DID WE REALLY TALK TO AI?

Experience Sharing of CKIP at
Talk-to-AI Grand Challenge

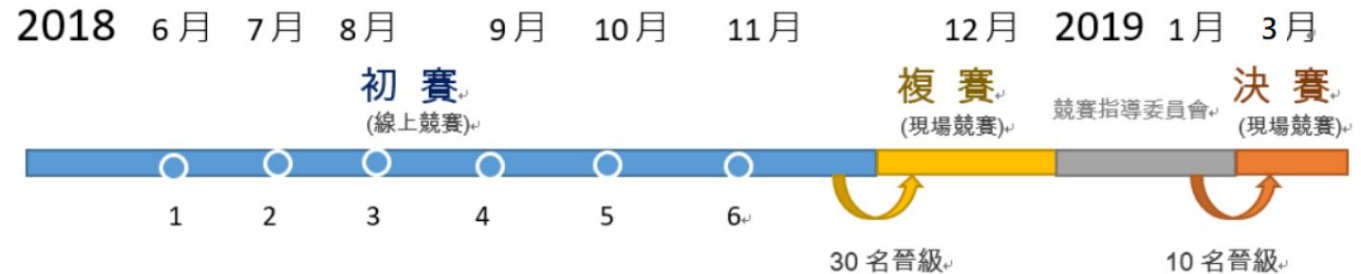
2019-08-16

賽 程

- 初賽
 - ◆ 選擇題
 - ◆ 線上競賽 (Kaggle 上傳)

- 複賽
 - ◆ 選擇題
 - ◆ 現場競賽

- 決賽
 - ◆ 選擇題及簡答題
 - ◆ 現場競賽



任務－選擇題

- Sound files
 - ◆ A. context
 - ◆ B. question
 - ◆ C. options (short-answer questions do not have this part)
- E.g. (with ASR error)
 - ◆ A: 明年元旦起，交通部新規定，汽車輪胎胎紋深度將納入定期檢驗項目之一，一段，深度未達1.6公里，近一個月內沒換胎，將會被吊銷牌照，民衆除了定期檢驗臺文，也可以自己用十元硬幣檢測，只要看得見國父像一頂下緣，表示該換輪胎了。
 - ◆ B: 汽車胎紋未達到多少公里將會被吊銷牌照？
 - ◆ C: 11.5公里，21.4公里，31.3公里是1.6公里。

必需要處理：C選項切分

- ASR problem
 - ◆ The option part comes in one sound file without a clear split for 4 options
 - ◆ A [parse algorithm] is needed
- [Parse algorithm]
 - ◆ Input: option text segments provided by ASR
 - E.g. [“一起勿！”，“二晴天；三雨天四陰天。”]。
 - ◆ Output: 4 text segments
 - E.g. [“起勿”，“晴天”，“雨天”，“陰天”]

任務－簡答題

- Sound files

- ◆ A. context
- ◆ B. question

- E.g.

- ◆ A: 請聽這段話，然後回答問題臺東聖母醫院是個小型醫院，除了醫治病人，還肩負起照顧臺東弱勢原住民的重責大任！他的廚房每天提供午餐給獨居老人和兒童，讓他們不至於捱餓但7月份的強烈颱風重創臺東，聖母醫院的倉庫被淹沒，導致百米泡水，連供應午餐的廚房設備也嚴重毀損，醫院估計重建經費逼近千萬聖母醫院表示，他們每天用的食材一向靠志工栽種的有機蔬果供應此外醫院向政府承租農地，一年兩戶的稻米還可以小包裝透過網路販售，扣除成本之後，販賣的銀魚，就是維持農場漢！廚房營運的資金然而風災過後收成全無，設備毀損，復原之路十分漫長！聖母醫院！因而呼籲各界人士慷慨解囊！協助早日重建，讓他們的廚房不至於斷炊。
- ◆ B: 醫院的食物來自哪裏？

RULE-BASED METHODS (FEATURE ENGINEERING)

語料

● 使用的實驗語料來源

◆ 華語文能力測驗 TOCFL

◆ <https://www.sc-top.org.tw/mocktest.php>

◆ 挑選320題

- 1 | 1. ↓
- 2 | 男：媽，你看見我的課本嗎？ ↓
- 3 | 女：書桌那裡沒有嗎？ ↓
- 4 | 男：找過了，都沒有。等一下就要上課了，真糟糕。 ↓
- 5 | 女：不是放在書桌上，就是跟字典放在一起，你去看看。 ↓
- 6 | 男：這位太太的意思是什麼？ ↓
- 7 | (A) 她要兒子再找找 (B) 課本不在書桌上 (C) 她把課本收起來了 (D) 書桌上放了本字典 ↓
- 8 | answer: A ↓
- 9 | ↓
- 10 | 2. ↓
- 11 | 男：好久沒來這裡了，變了好多喔。 ↓
- 12 | 女：以前這裡是漂亮的風景，現在這裡到處都是人。 ↓
- 13 | 男：人多不好嗎？ ↓
- 14 | 女：我喜歡安靜的環境。 ↓
- 15 | 男：這位小姐最後的意思是什麼？ ↓
- 16 | (A) 這裡很吵 (B) 這裡人不多 (C) 這裡的環境很好 (D) 這裡的風景很美 ↓
- 17 | answer: A ↓
- 18 | ↓

範例題目

Context:

男 喂，媽，請寄一些錢給我，我要買衣服。

女 衣服不是都帶去了嗎？為什麼還要買？

男 我變胖了，那些衣服都穿不下了。

女 你才去念了半年的書就變胖了，我想那裡的東西一定很好吃。

Question:

男 這個兒子打電話跟母親說什麼？

Options:

- (A) 沒收到衣服
- (B) 沒錢吃東西
- (C) 沒錢買衣服
- (D) 請母親寄衣服

範例題目

Context:

男 喂，媽，請寄一些錢給我，我要買衣服。

女 衣服不是都帶去了嗎？為什麼還要買？

男 我變胖了，那些衣服都穿不下了。

女 你才去念了半年的書就變胖了，我想那裡的東西一定很好吃。

Question:

男 這個兒子打電話跟母親說什麼？

Options:

(A) 沒收到衣服

(B) 沒錢吃東西

(C) 沒錢買衣服

(D) 請母親寄衣服

← Similarity Comparison seems good to solve it.

範例題目

Context:

男 請進，有什麼事嗎？

女 老闆，我最近總是覺得很累、很不舒服。我想請假。

男 妳想請多久呢？

女 如果可以的話，我想請一個月；但是如果不行，兩個星期也可以。

Question:

男 這位小姐希望老闆答應她什麼事？

Options:

- (A) 讓她休息一個月
- (B) 讓她的工作少一點
- (C) 兩星期後，可以休息幾天
- (D) 一個月後，讓她離開公司



Similarity Comparison is not enough. Either A or D is possible

範例題目

Context:

男 請進，有什麼事嗎？

女 老闆，我最近總是覺得很累、很不舒服。我想請假。

男 妳想請多久呢？

女 如果可以的話，我想請一個月；但是如果不行，兩個星期也可以。

Question:

男 這位小姐希望老闆答應她什麼事？

Options:

- (A) 讓她休息一個月
- (B) 讓她的工作少一點
- (C) 兩星期後，可以休息幾天
- (D) 一個月後，讓她離開公司



Common senses tell us the following possible cause-effect reasoning:

累->休息

請假->休息

廣義知網下的詞彙定義與推論

BeBad|衰變

tired|疲乏 [大困馬乏, 乏力, 困乏, 困阨, 困倦, 困竭, 委頓, 倦, 倦怠, 疲, 疲乏, 疲困, 疲於奔命, 疲倦, 疲累, 疲軟, 疲勞, 疲頓, 疲置, 疲憊, 累, 勞乏, 勞累, 勞瘁, 勞頓, 無神, 暈, 困乏, 萎頓, 瘁, 慵, 憊, 僵, 羸憊, 勞, 乏, 困頓, 暈倦]

【塵勞】 [舟車勞頓, 塵勞]

【精疲力竭】 [爛額, 筋疲力盡]

Miscellaneous

HungryThirsty| [食不充饑, 挨餓, 嗷嗷待哺, 餓, 餓]

ill|病態 [害, 不]

義原訊息	
Primitive:	tired 疲乏
定義式:	{BeBad 衰變:cause={exhaust 損耗}}
Event frame:	THEME{theme}
Primitive relations:	possible consequence={rest 休息}
Conflation of events:	theme({tired 疲乏})=agent({rest 休息})

焦頭, 空腹, 渴, 病倒

cease|停做 [中輟, 中斷, 叫停, 打住, 休, 收, 收手, 收尾, 收腳, 住手, 拉倒, 洗手不幹, 停止]

GiveUp|戒除 [戒, 戒除, 戒掉, 改掉, 忌, 斷]

rest|休息 [喘氣, 休息, 休歇, 休養生息, 休憩, 將息, 歇息, 解裝, 緩氣, 憩, 憩息, 喘息]

sleep|睡 [昏睡, 眠, 寐, 就寢, 睡, 睡下, 睡覺, 蒙頭大睡, 休眠, 休眠]

【小憩】 [小憩, 助, 會, 斷助]

累->休息
請假->休息

詞彙訊息	
詞彙:	請假
詞性:	VA4
英文意涵:	ask for leave
Event Frame:	
Primitive relations:	
Conflation of events:	
定義式: [編輯]	{request 要求:content={enjoy 享受:content={cease 停做},purpose={rest 休息}}}

作答方式

- 隨機
- 挑最短題目
- 挑最長題目
- 詞彙比對
- 字/音比對
- 詞彙語意 EHowNet 比對
- EHowNet 推論資訊 比對
- Word Association 比對
- 問題類型

KAGGLE3

• 上傳的分數

#	Δ1w	Team Name	Score 📊	Entries
1	new	hungyilee	0.62600	5
2	new	Intellection	0.61333	5
3	new	水筆仔隊	0.59533	5
4	new	Alman	0.58066	5
5	new	The MUG	0.57933	5
6	new	d204096001	0.55000	5
7	new	m3upg6u04g4	0.53333	5
8	new	Wei-Yun Ma	0.53266	5

• 目前總排名

User Name	Kaggle1	Kaggle2	Kaggle3	總計
hungyilee	20	16	20	56
aieverywhere	14	18	16	48
kandelia	14	16	18	48
intellection	8	20	18	46
m3upg6u04g4	18	10	14	42
peichieh	18	12	10	40
ctcasus	14	14	12	40
chesterkuo	14	12	12	38
elandtw	12	14	10	36
hakka1	16	10	8	34
d204096001	0	18	14	32
aiedward	16	8	6	30
consciouswithlife	8	10	10	28
hsinminwang	2	12	14	28
huntfoxsu	0	14	12	26
weiyunma	0	12	14	26
chhsiao	12	6	4	22
stegben	12	6	2	20
yifuchen0	10	6	4	20

OPTION-STORY FUZZY MATCH

- Consider whether words in option also occur in story
- Consider whether characters in option also occur in story
- Consider whether phonemes in option also occur in story
- Consider whether word-word dependency in option also occur in story

HIERARCHICAL N-GRAM MATCH

題目/Story

詞

 $SW_1 SW_2 SW_3 SW_4 SW_5$ SW_m

問題/Question

詞

 $QW_1 QW_2 QW_3 \dots$

選項1/Option1

選項2/Option2

選項3/Option3

選項4/Option4

詞

 $W_1 W_2 W_3 \dots$ $W_1 W_2 W_3 \dots$ $W_1 W_2 W_3 \dots$ $W_1 W_2 W_3 \dots$

字

 $c_{11} c_{12} c_{21} \dots$

音

 $f_{11} f_{12} f_{21} \dots$

HIERARCHICAL N-GRAM MATCH

$$\text{range} = \underset{\text{continue_3_sents}}{\operatorname{argmax}} \text{match}(\text{question}, \text{continue_3_sents})$$

$$\text{answer} = \max_{\text{option}} \text{match}(\text{option}, \text{range})$$

$$\text{match}(\mathbf{A}, \mathbf{B}) = \sum_{\text{scope} \in \mathbf{A}} \text{match_score}(\text{scope}, \mathbf{B})$$

match order ↓

match_score	scope	next pointer
8 if scope $\in B$	$W_i W_{i+1} W_{i+2} W_{i+3}$	W_{i+4}
5 if scope $\in B$	$W_i W_{i+1} W_{i+2}$	W_{i+3}
3 if scope $\in B$	$W_i W_{i+1}$	W_{i+2}
1 if scope $\in B$	W_i	W_{i+1}
0.5 if scope $\in B$	$C_1(W_i) C_2(W_i) \dots$	W_{i+1}
0.25 if scope $\in B$	$f(C_1(W_i)),$ $f(C_2(W_i)) \dots$	W_{i+1}
0 otherwise		W_{i+1}

HIERARCHICAL N-GRAM MATCH

● 參考資源

◆ 同音 / 近音同調字

漢字	同音同調	近音同調
人	仁王兒任	瓢攘仍襁
麼	邀要啣夭妖吆腰	幽攸悠優憂
亢	伉炕抗	艮互楨
...		
共有 5401 字 [Liu et al., 2011]		

● 判斷 問題 是否為 反向 選擇？

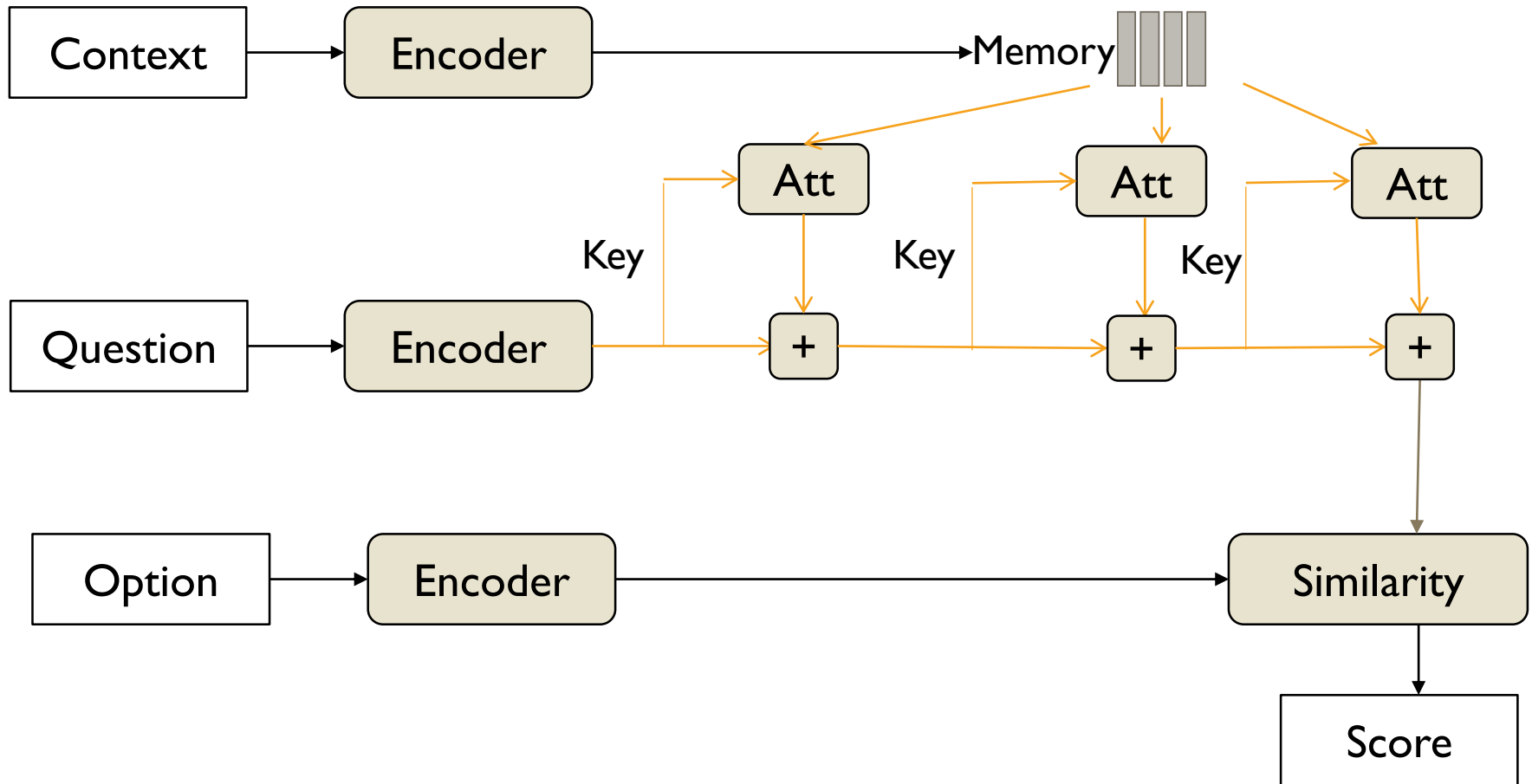
- ◆ 參考詞彙：不是，不包含，不能夠，不對，...，等共有 20個

HIERARCHICAL N-GRAM MATCH

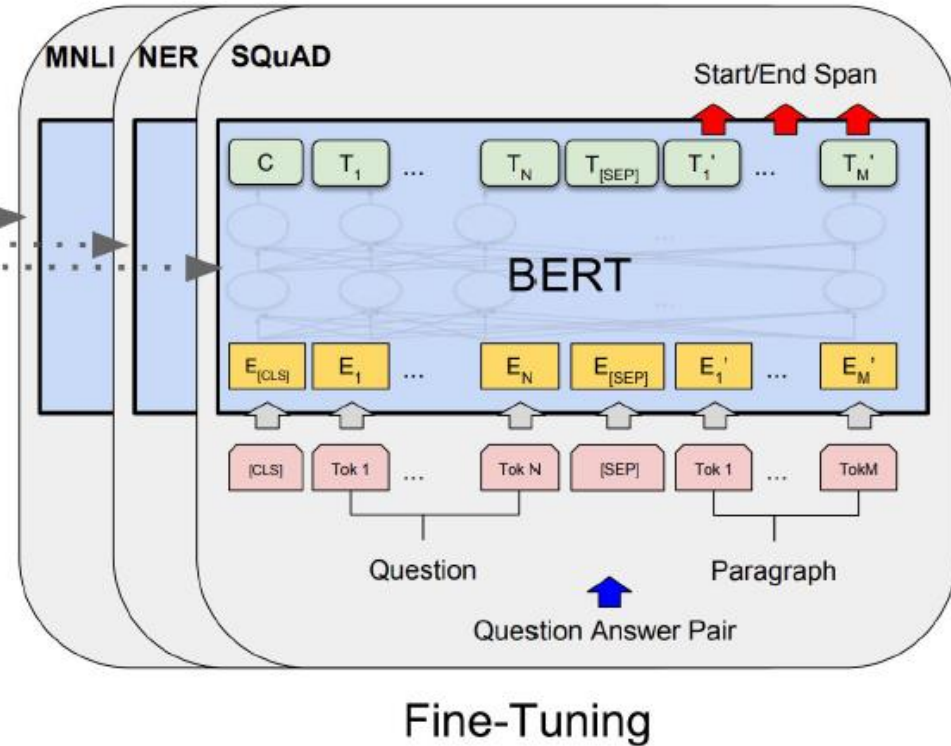
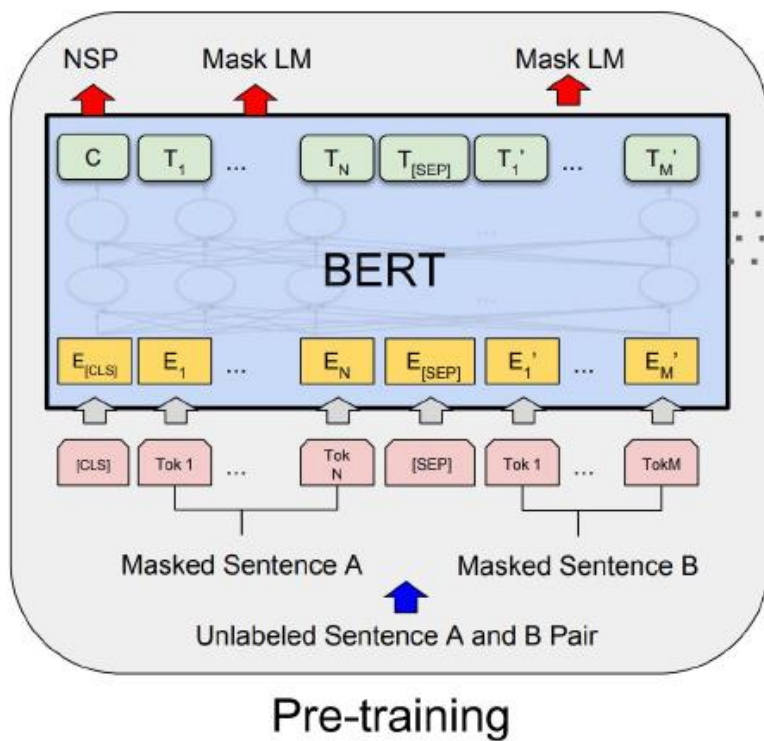
- Option-Context Matching
 - ◆ The window size of candidates: three sentences
 - ◆ Candidates match score
 - Basic score: question match
 - Match score for each option
 - Algorithm(n-gram)
 - word token → character string → pinyin token → pinyin string
 - ◆ Choose the option with the max/min(negative question) score
- Validate
 - ◆ Semi-Final-2-ASR
 - ◆ Accuracy: 0.635

DATA-DRIVEN METHODS

MULTI-HOP ATTENTION QA



BERT



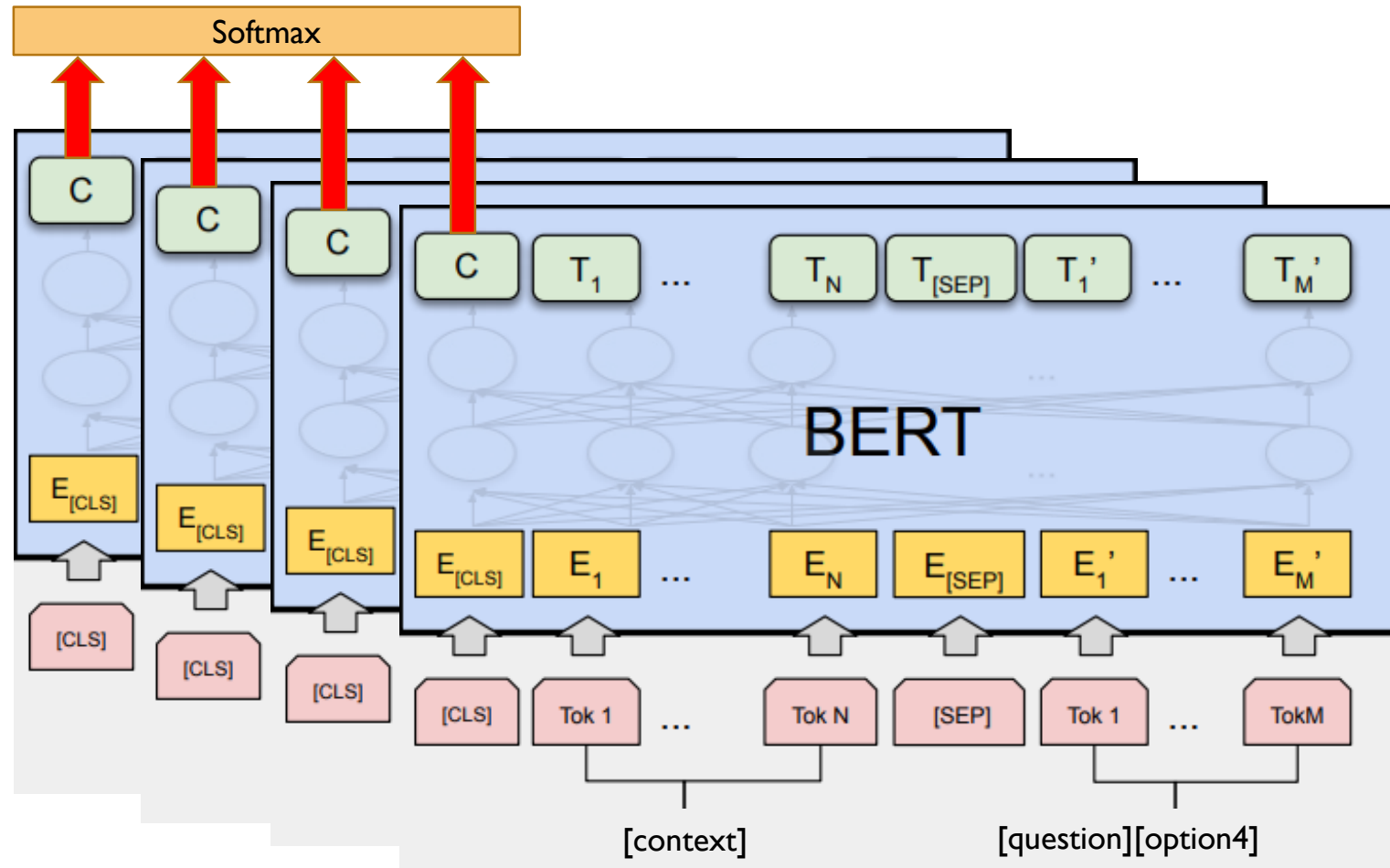
BERT-SQUAD OPTION SCORING

- Input:
 - generated answer, e.g. “霧天”
 - an option, e.g. “起勿”
- Output: score $\in [0,1]$
 1. For each character
 - ◆ Transform to the representative word of its “sound synset”.
 - Sound synset, e.g. {兀戌晤惡寤驚物勿悟塢霧...}
 - Transform, e.g. “霧天” -> “兀天”, “起勿” -> “起兀”
 2. Compute sequence matching ratio
 - ◆ Matching_length / average length
 - E.g. $\text{len}(\text{“兀”}) / \text{avglen}(\text{“兀天”, “起兀”}) = 1/2$

BERT-SQUAD

- Generative
- Train
 - ◆ DRCD
- Validate
 - ◆ Semi-Final-2-ASR
 - ◆ Accuracy: 0.705

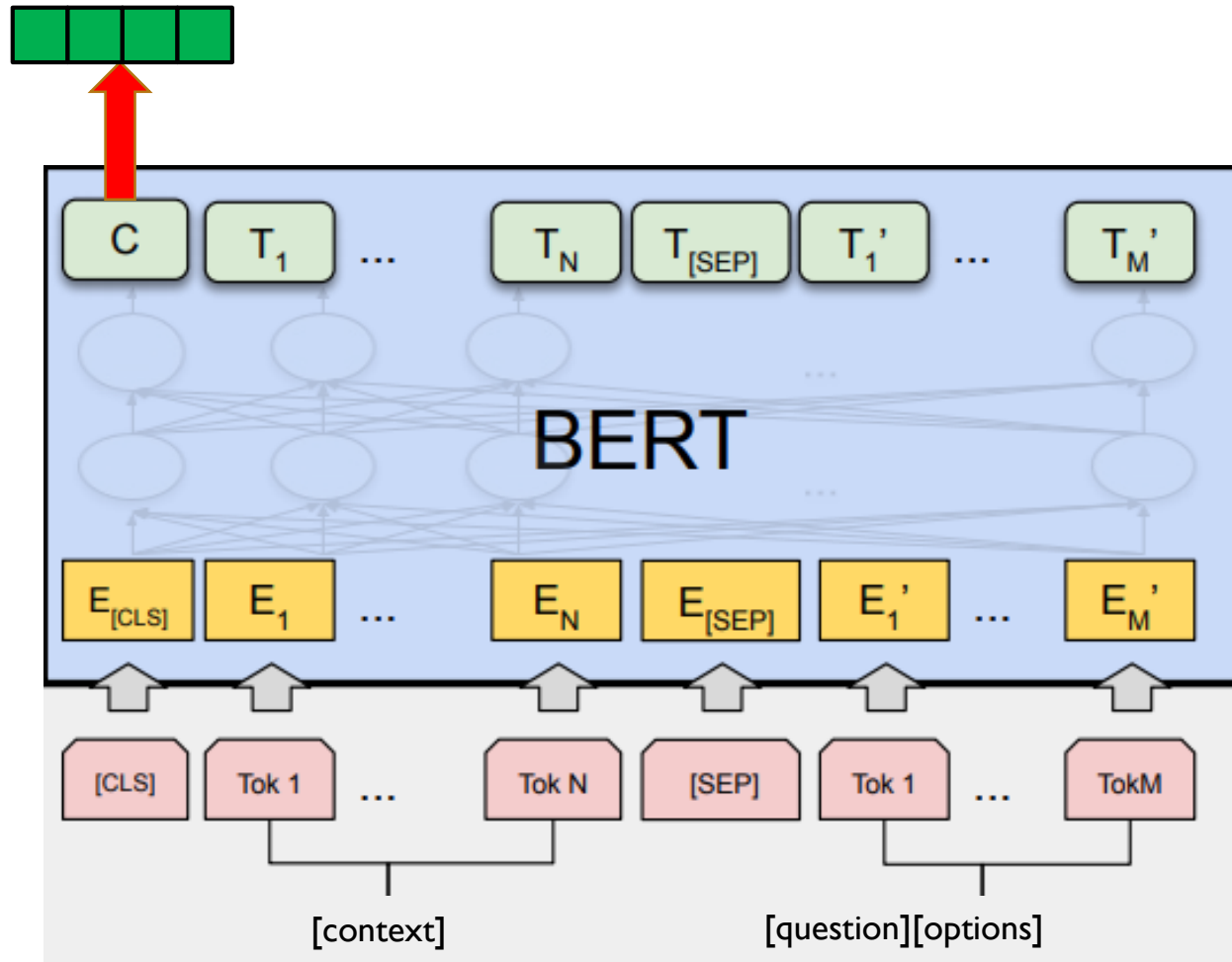
BERT-INDEPENDENT



BERT-INDEPENDENT

- Train
 - ◆ Kaggle-123456-ASR
 - ◆ Semi-Final-1-ASR
- Validate
 - ◆ Semi-Final-2-ASR
 - ◆ Accuracy: 0.759

BERT-CONCATENATE



BERT-CONCATENATE

- Train
 - ◆ Kaggle-123456-ASR
 - ◆ Kaggle-123456-transcript
 - ◆ Semi-Final-1-ASR
 - ◆ Semi-Final-1-transcript
 - ◆ DRCD-multiple-choice
- Validate
 - ◆ Semi-Final-2-ASR
 - ◆ Accuracy: 0.773

CREATE DRCD-MULTIPLE-CHOICE DATASET

- Extract similar POS/NER sequences in DRCD context as options
 1. Ground truth: 2007年1月16日['Nd', 'Nd', 'Nd'] (from DRCD training set, id: 1001-10-2)
 2. ['Nd', 'Nd', 'Nd']
 - E.g. 2004年8月5日
 3. ['Nd', 'Nd']
 - E.g. 2014年6月
 4. ['Nd']
 - E.g. 2010年, 2023年

ENSEMBLE

COMBINATIONS

- Grid-search best weights $\in (0,1,2,\dots,9)$

Option-Context Matching	BERT-squad	BERT-independent	BERT-concatenate	Semi-Final-2	Final
1				0.635	
	1			0.705	
		1		0.759	
			1	0.773	
		2	3	0.790	
	2	1	1	0.816	
9	9	6	7	0.821	0.475

- Better models need not have higher weights

總成績第一名		總成績第二名	
選擇題成績(正確率)	簡答題成績(正確率)	選擇題成績(正確率)	簡答題成績(正確率)
0.537	0.16	0.478	0.08

名次	團隊名稱	選擇題成績(正確率)	簡答題成績(正確率)
3	團隊8 weiyuma	0.475	0.28
4	團隊3 elandtw	0.432	0.08
5	團隊2 d204096001	0.422	0.08
6	團隊1 aieverywhere	0.408	0.16
7	團隊7 THE MUG	0.359	0.32
8	團隊4 hakka1	0.320	0.16

ASR + SYSTEM

SPEECH RECOGNITION

- Google Cloud Platform ASR
- Olami ASR
- iFlytek ASR
 - ◆ 語音聽寫 (IAT)
 - ◆ 語音轉寫 (LFASR)

SPEECH RECOGNITION – GOOGLE CLOUD

- Google Cloud Platform ASR
 - ◆ Python API
 - ◆ Pros:
 - Easy to use
 - ◆ Cons:
 - No Chinese punctuation
 - Daily/Hourly task limitation

SPEECH RECOGNITION – OLAMI ASR

- Olami ASR – 語音聽寫 (IAT)
 - ◆ REST API
 - ◆ Pros:
 - Fast
 - ◆ Cons:
 - Bad performance
 - 1 minute limitation
 - Daily task limitation (About 500 tasks)

SPEECH RECOGNITION – IFLYTEK

ASR – IAT

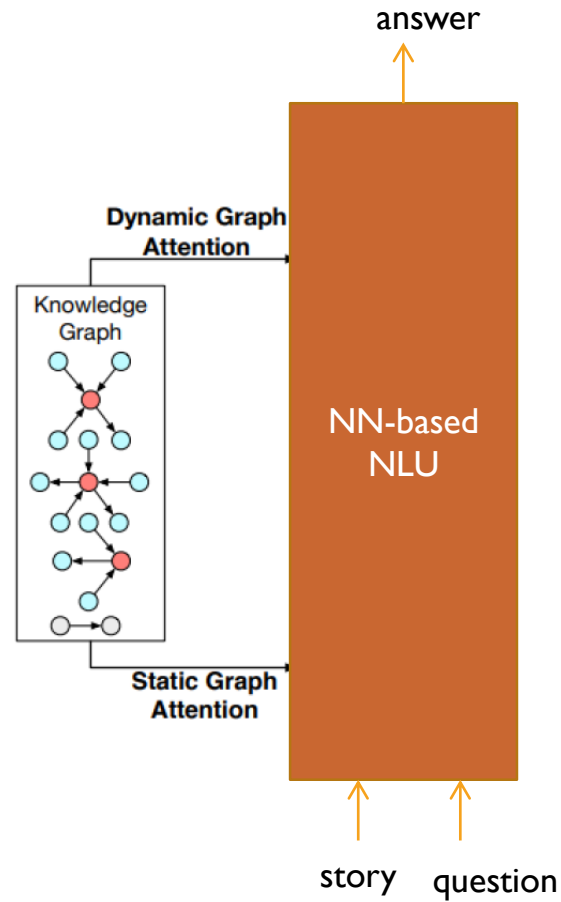
- iFlytek ASR – 語音聽寫 (IAT)
 - ◆ REST API
 - ◆ Pros:
 - Fast
 - ◆ Cons:
 - Bad performance
 - 1 minute limitation
 - Daily task limitation (Undefined)

SPEECH RECOGNITION – IFLYTEK ASR – LFASR

- iFlytek ASR – 語音轉寫 (LFASR)
 - ◆ JAVA API
 - ◆ Pros:
 - Great performance
 - No limitation on wave
 - No limitation on task
 - ◆ Cons:
 - slow

DID WE REALLY TALK TO AI?

POSSIBLE SOLUTION



THANKS FOR YOUR KINDLY
ATTENTION