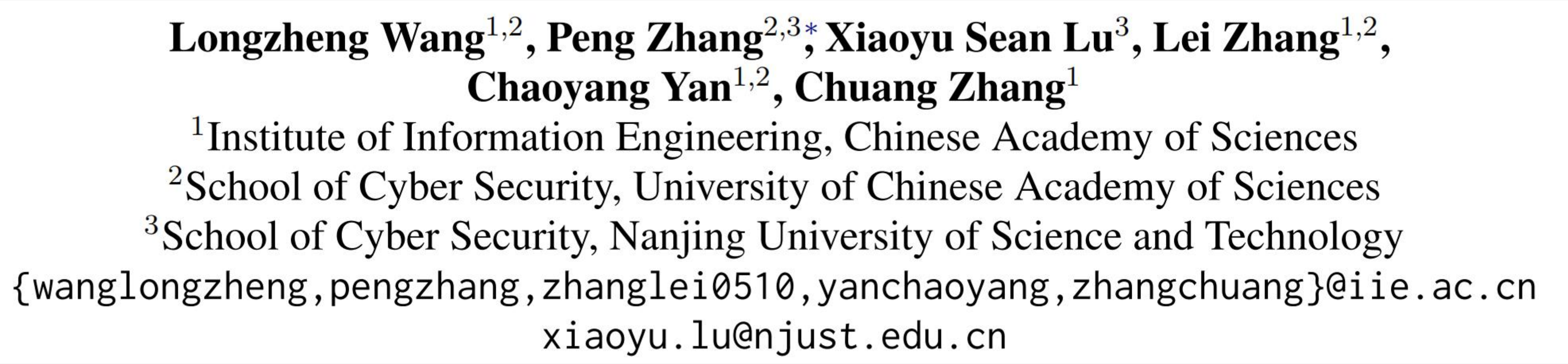
2022_EMNLP_QaDialMoE Question-answering Dialogue based Fact Verification with Mixture of Experts Tracking



Longzheng Wang[1,2], Peng Zhang[2,3]*, Xiaoyu Sean Lu[3], Lei Zhang[1,2], Chaoyang Yan[1,2], Chuang Zhang[1]

[1]Institute of Information Engineering, Chinese Academy of Sciences
[2]School of Cyber Security, University of Chinese Academy of Sciences
[3]School of Cyber Security, Nanjing University of Science and Technology
{wanglongzheng,pengzhang,zhanglei0510,yanchaoyang,zhangchuang}@iie.ac.cn
xiaoyu.lu@njust.edu.cn

—— *EMNLP 2022*

https://github.com/wishever/QaDialMoE

Reported by Yuyang Lai

1.Introduction

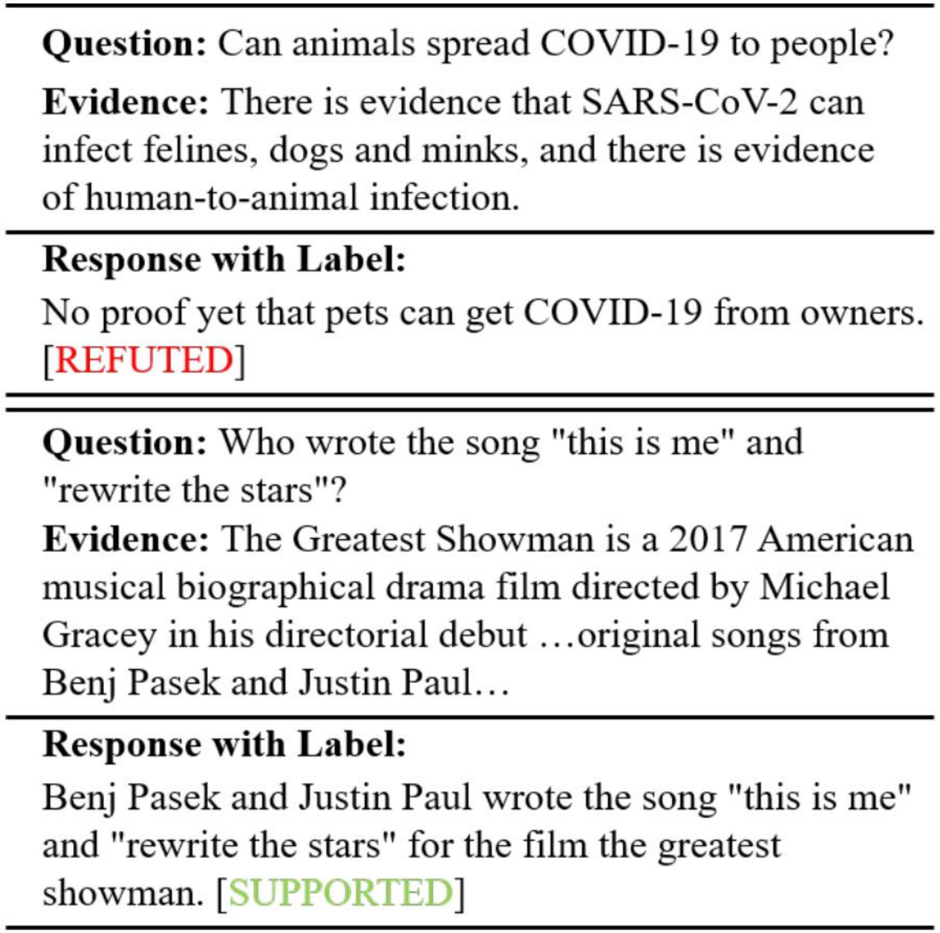
2.Method

3.Experiments

Introduction

Previous works for the fact verification mainly focused on reasoning against pieces of evidence from Wikipedia passages, while rarely considered questions sought by Internet users.

However, the questions also contain rich information to support the fact verification.



Question: Can animals spread COVID-19 to people?

Evidence: There is evidence that SARS-CoV-2 can infect felines, dogs and minks, and there is evidence of human-to-animal infection.

Response with Label:

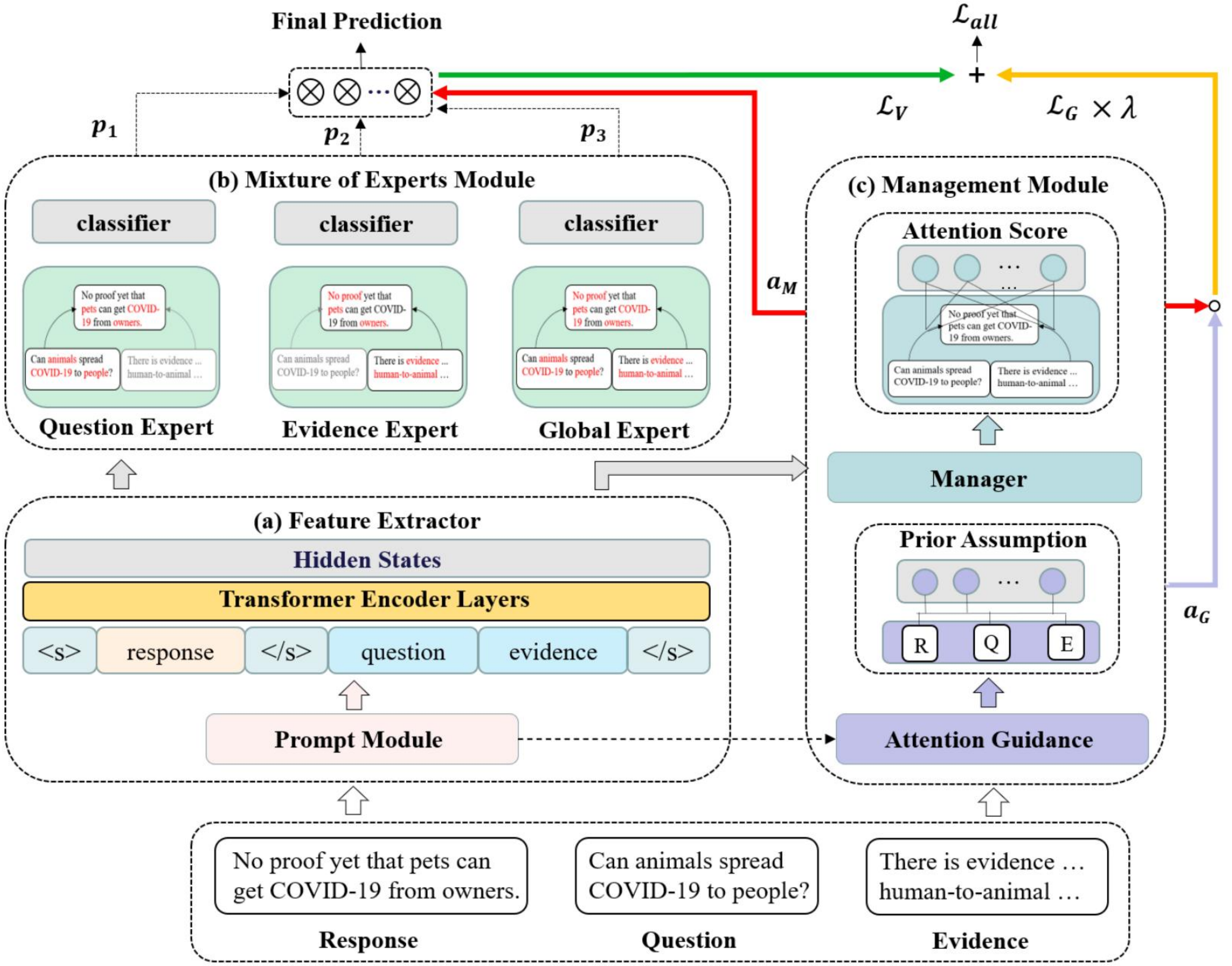
No proof yet that pets can get COVID-19 from owners. [REFUTED]

Question: Who wrote the song "this is me" and "rewrite the stars"?

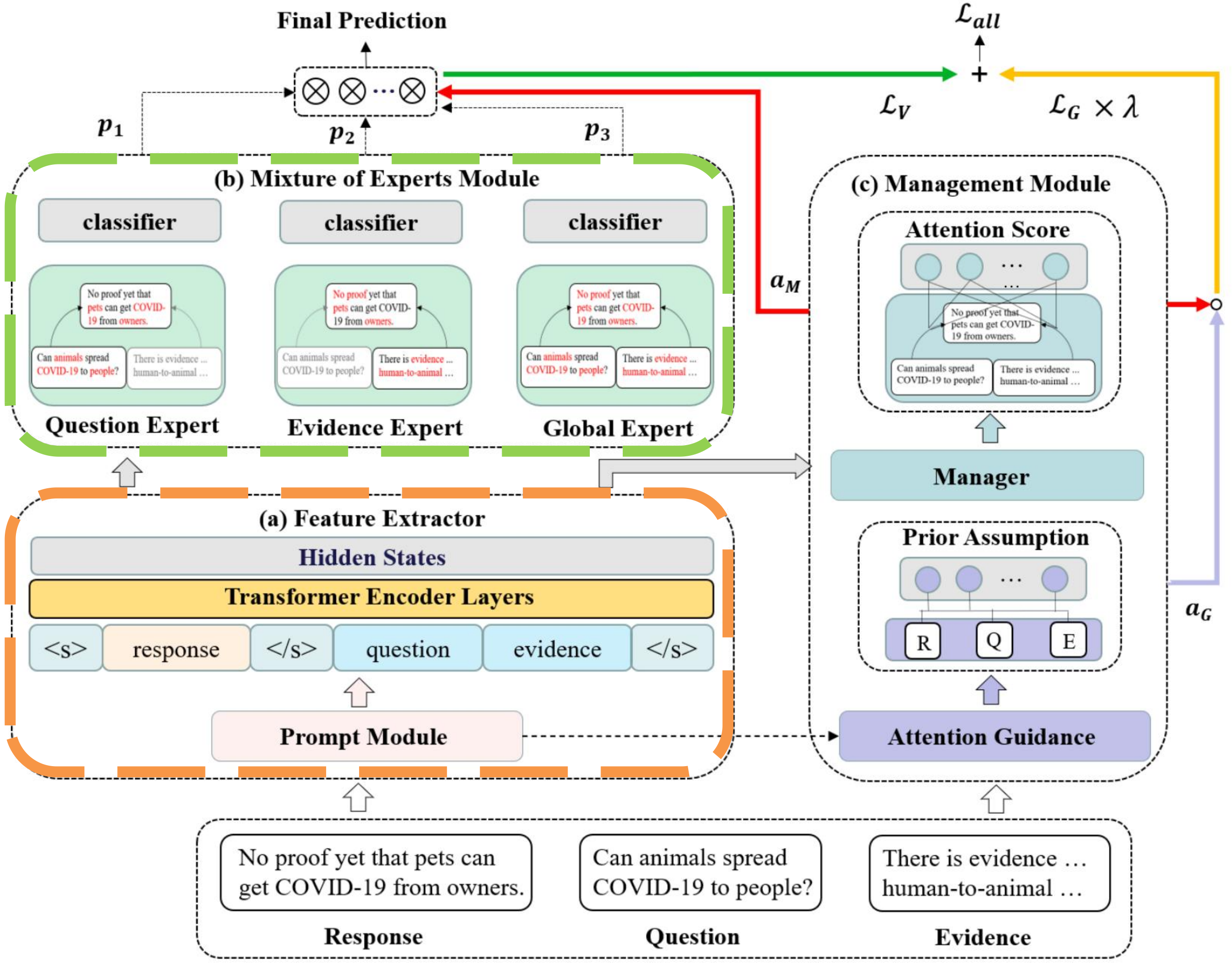
Evidence: The Greatest Showman is a 2017 American musical biographical drama film directed by Michael Gracey in his directorial debut …original songs from Benj Pasek and Justin Paul…

Response with Label:

Benj Pasek and Justin Paul wrote the song "this is me" and "rewrite the stars" for the film the greatest showman. [SUPPORTED]


Final Prediction
$\mathcal{L}_{all}$
+
$\mathcal{L}_V$
$\mathcal{L}_G \times \lambda$
p_1
p_2
p_3
a_M
a_G
(b) Mixture of Experts Module
classifier
classifier
classifier
No proof yet that pets can get COVID-19 from owners.
Can animals spread COVID-19 to people?
There is evidence ... human-to-animal ...
Question Expert
Evidence Expert
Global Expert
(c) Management Module
Attention Score
Manager
Prior Assumption
R
Q
E
Attention Guidance
(a) Feature Extractor
Hidden States
Transformer Encoder Layers
<s>
response
</s>
question
evidence
</s>
Prompt Module
No proof yet that pets can get COVID-19 from owners.
Can animals spread COVID-19 to people?
There is evidence … human-to-animal …
Response
Question
Evidence

Method


Final Prediction
$\mathcal{L}_{all}$
+
$\mathcal{L}_V$
$\mathcal{L}_G \times \lambda$
p_1
p_2
p_3
a_M
a_G
(b) Mixture of Experts Module
classifier
classifier
classifier
No proof yet that pets can get COVID-19 from owners.
Can animals spread COVID-19 to people?
There is evidence ... human-to-animal ...
Question Expert
Evidence Expert
Global Expert
(c) Management Module
Attention Score
Manager
Prior Assumption
R
Q
E
Attention Guidance
(a) Feature Extractor
Hidden States
Transformer Encoder Layers
<s>
response
</s>
question
evidence
</s>
Prompt Module
No proof yet that pets can get COVID-19 from owners.
Can animals spread COVID-19 to people?
There is evidence … human-to-animal …
Response
Question
Evidence

Feature Extractor

$$\mathbf{H} = f_{LM}(\mathbf{L}_{r,q,e}) \qquad (1)$$

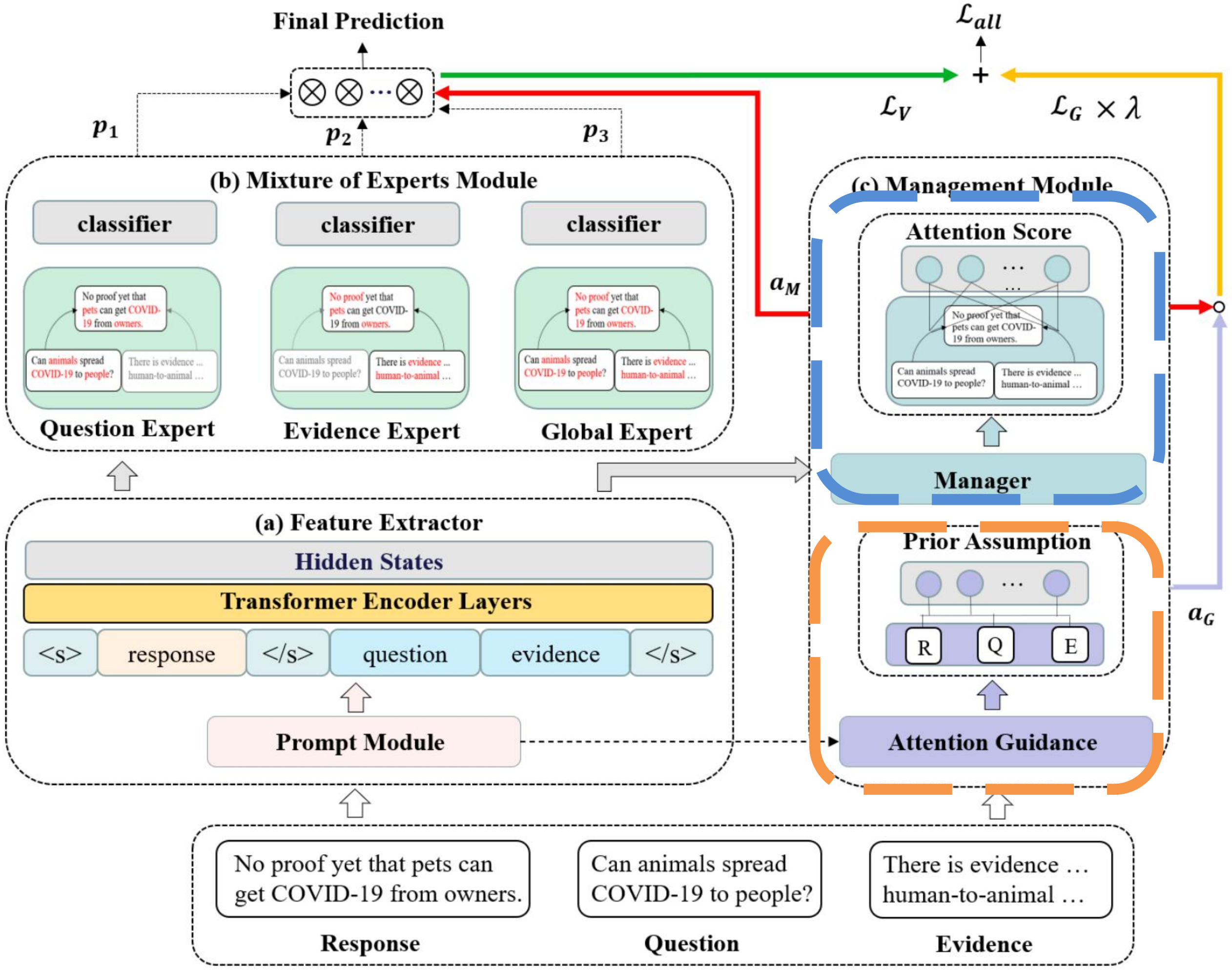
$$\mathbf{L}_{r,q,e} = [\langle s\rangle, \mathbf{R}, \langle/s\rangle, \mathbf{Q}, \mathbf{E}, \langle/s\rangle]$$

Mixture of Experts Module

$$\mathbf{h}_i = f_{Enc_i}(\mathbf{H}) \qquad (2)$$

$$\mathbf{p}_i = softmax(tanh(\mathbf{h}_i\mathbf{W}_1^i)\mathbf{W}_2^i) \qquad (3)$$

Method



Attention Guidance

response-question pair, response-evidence pair

response-question-evidence pair

$$\mathbf{z}_0 = ((\mathbf{z}_0)_0, (\mathbf{z}_0)_1, (\mathbf{z}_0)_2)^T = (0.2, 0.2, 0.6)^T$$

$$\delta_i = a_i(1 - s_i)^2$$

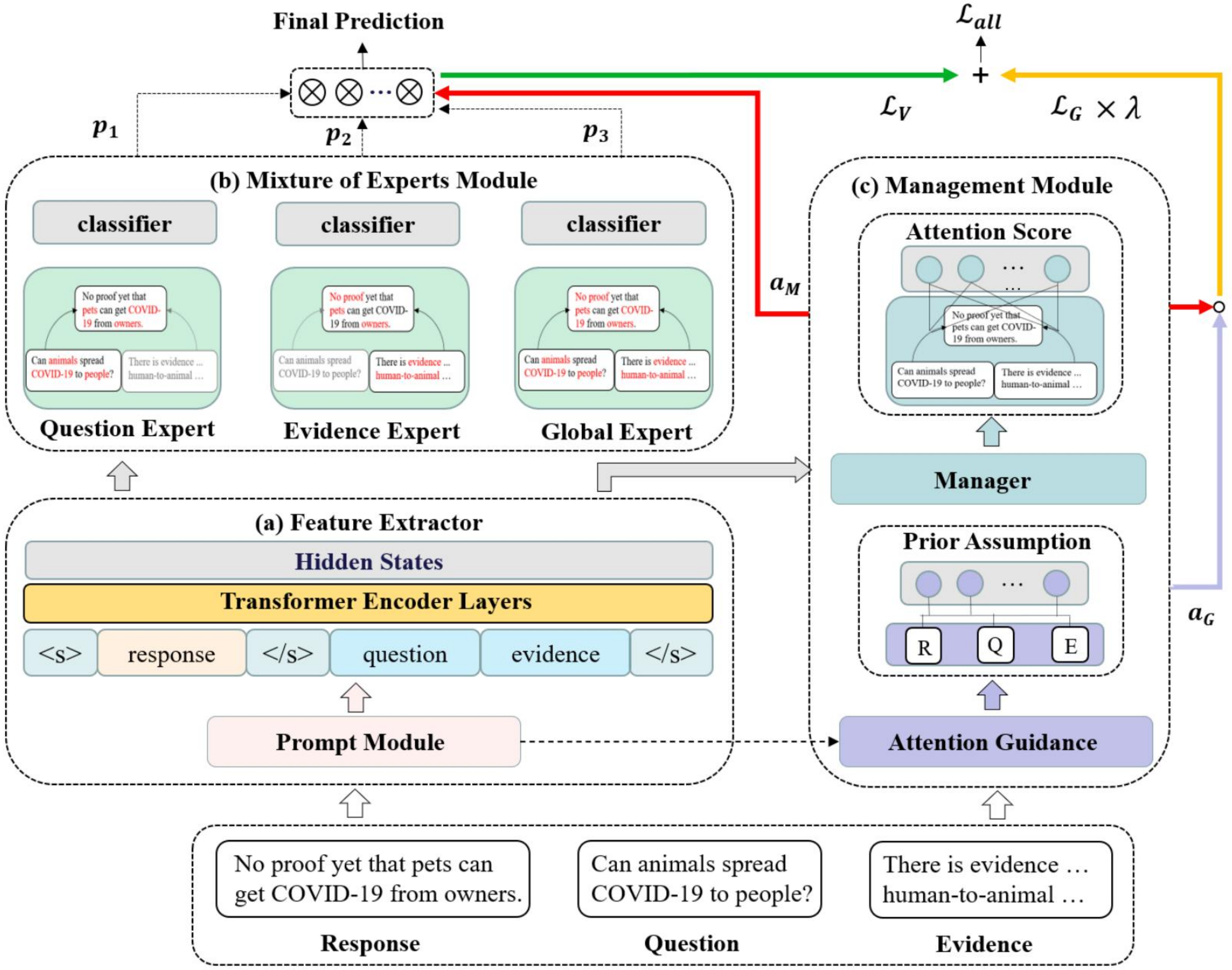
$$\mathbf{a}_G = softmax(\mathbf{z}_0 + \delta)$$

Manager

$$\mathbf{h}_M = f_{Enc_M}(\mathbf{H}) \qquad (4)$$

$$\mathbf{a}_M = softmax(tanh(\mathbf{h}_M \mathbf{W}_1^M)\mathbf{W}_2^M) \qquad (5)$$

Method



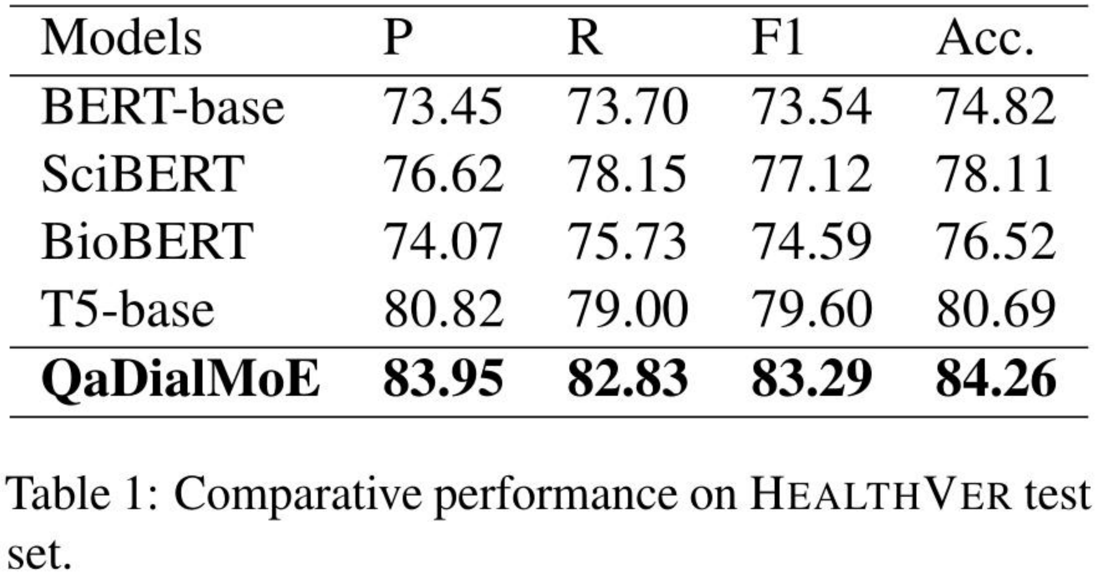
Verification Loss

$$\mathcal{L}_V = \sum_{i=1}^{n_e} (\mathbf{a}_M)_i \cdot H_{CE}(\mathbf{p}_i, l) \qquad (6)$$

Guidance Loss

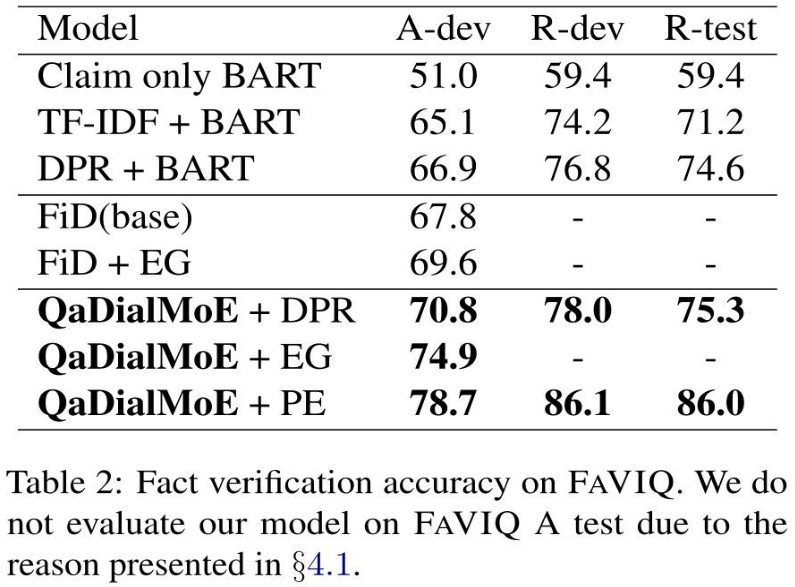
$$\mathcal{L}_G = D_{KL}(\mathbf{a}_G || \mathbf{a}_M) \qquad (7)$$

Experiments



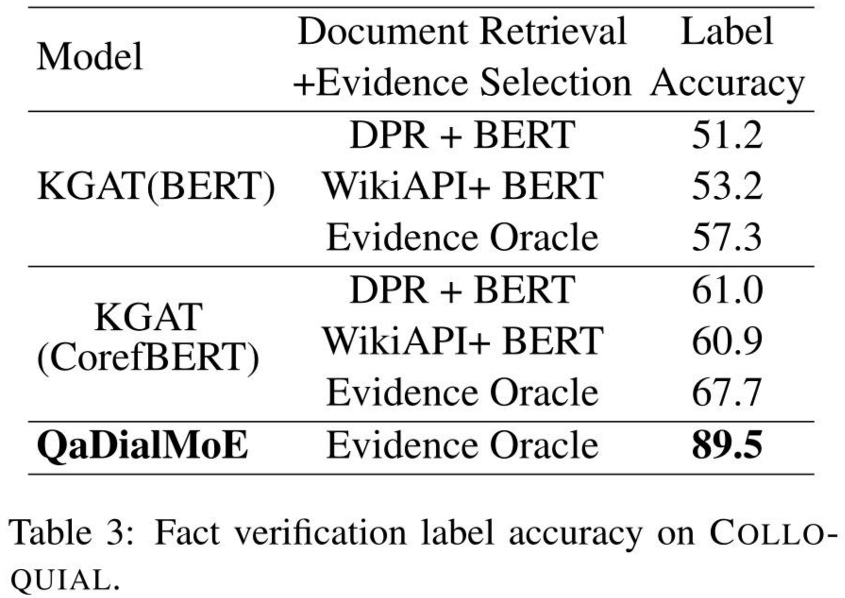
Models	P	R	F1	Acc.
BERT-base	73.45	73.70	73.54	74.82
SciBERT	76.62	78.15	77.12	78.11
BioBERT	74.07	75.73	74.59	76.52
T5-base	80.82	79.00	79.60	80.69
QaDialMoE	**83.95**	**82.83**	**83.29**	**84.26**

Table 1: Comparative performance on HEALTHVER test set.



Model	A-dev	R-dev	R-test
Claim only BART	51.0	59.4	59.4
TF-IDF + BART	65.1	74.2	71.2
DPR + BART	66.9	76.8	74.6
FiD(base)	67.8	-	-
FiD + EG	69.6	-	-
QaDialMoE + DPR	**70.8**	**78.0**	**75.3**
QaDialMoE + EG	**74.9**	-	-
QaDialMoE + PE	**78.7**	**86.1**	**86.0**

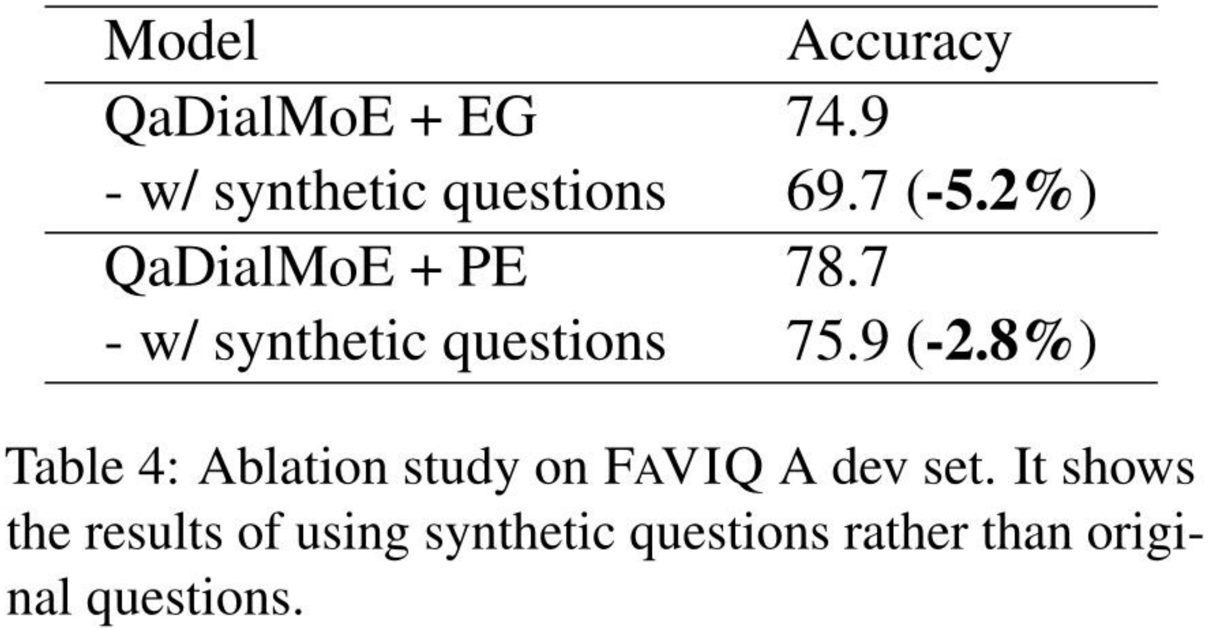
Table 2: Fact verification accuracy on FAVIQ. We do not evaluate our model on FAVIQ A test due to the reason presented in §4.1.



Model	Document Retrieval +Evidence Selection	Label Accuracy
KGAT(BERT)	DPR + BERT	51.2
	WikiAPI+ BERT	53.2
	Evidence Oracle	57.3
KGAT (CorefBERT)	DPR + BERT	61.0
	WikiAPI+ BERT	60.9
	Evidence Oracle	67.7
QaDialMoE	Evidence Oracle	**89.5**

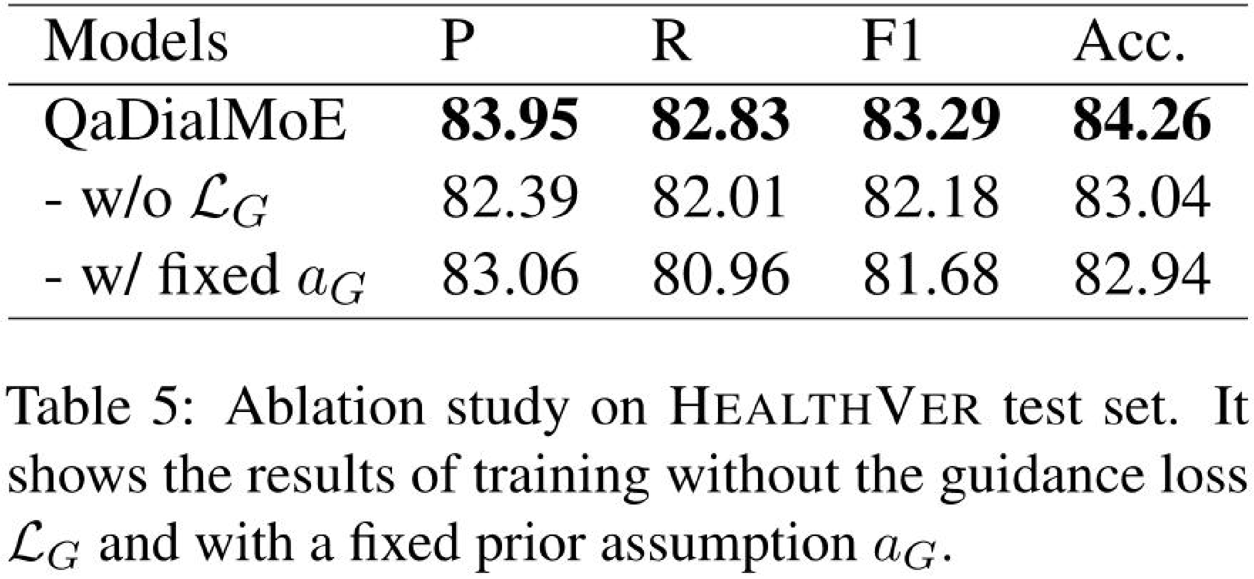
Table 3: Fact verification label accuracy on COLLOQUIAL.

Experiments



Model	Accuracy
QaDialMoE + EG	74.9
- w/ synthetic questions	69.7 (**-5.2%**)
QaDialMoE + PE	78.7
- w/ synthetic questions	75.9 (**-2.8%**)

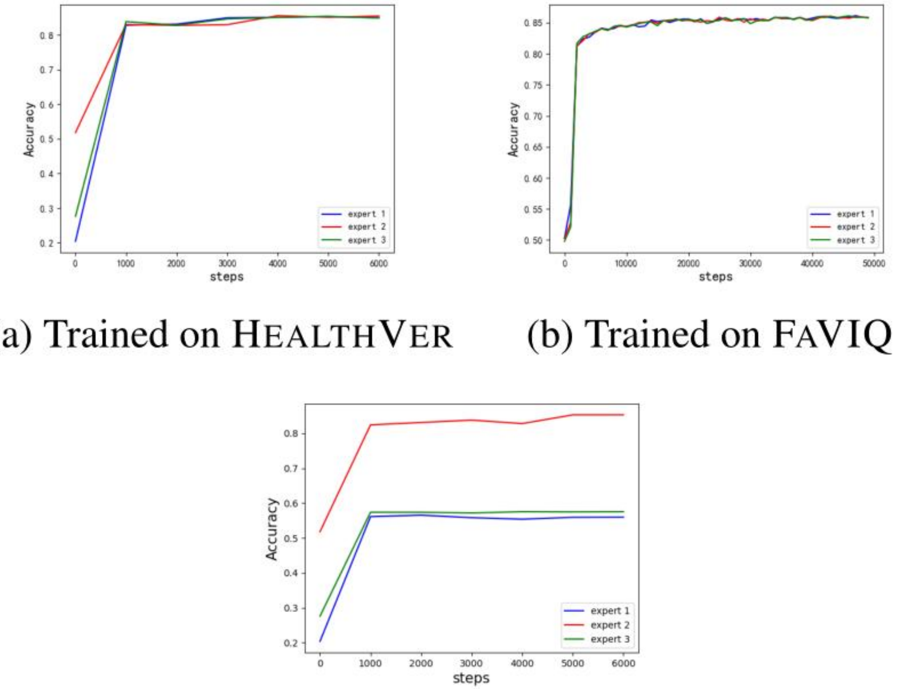
Table 4: Ablation study on FAVIQ A dev set. It shows the results of using synthetic questions rather than original questions.



Models	P	R	F1	Acc.
QaDialMoE	**83.95**	**82.83**	**83.29**	**84.26**
- w/o $\mathcal{L}_G$	82.39	82.01	82.18	83.04
- w/ fixed a_G	83.06	80.96	81.68	82.94

Table 5: Ablation study on HEALTHVER test set. It shows the results of training without the guidance loss $\mathcal{L}_G$ and with a fixed prior assumption a_G.

Experiments



(a) Trained on HEALTHVER (b) Trained on FAVIQ

(c) Trained on HEALTHVER without the Guidance Loss

Figure 3: The differentiation of experts. We show the model trained on HEALTHVER and FAVIQ R set with the *positive evidence*, and $n_e = 3$.

Thank you!