



## TCE at Qur'an QA 2022

Arabic Language Question Answering Over Holy Qur'an  
Using a Post-Processed Ensemble of BERT-based Models

---

Mohammed Elkomy, Moustafa F. Alzantot, Amany M. Sarhan

June 20, 2022

Tanta University

# Table Of Contents

1. Qur'an QA 2022 shared-task

2. Bert for QA

3. Span Voting Ensemble

4. Post-processing

Post-processing: Handling Sub-words

Post-processing: Redundancy Elimination

Uninformative Answer Removal

Updating the ranked list

Post-processing example

5. Official results

## Qur'an QA 2022 shared-task

---

## Bert for QA

---

We fine-tuned Bert for **extractive QA** [Devlin et al., 2019].

# Bert for QA

We fine-tuned Bert for **extractive QA** [Devlin et al., 2019].  
Two randomly initialized vectors  $S, E$  are stacked on top of Bert's representation.

# Bert for QA

We fine-tuned Bert for **extractive QA** [Devlin et al., 2019].  
Two randomly initialized vectors  $S, E$  are stacked on top of Bert's representation.

$$P_{start_i} = \frac{e^{S \cdot T_i}}{\sum_k e^{S \cdot T_k}}, P_{end_j} = \frac{e^{E \cdot T_j}}{\sum_k e^{E \cdot T_k}}$$

$$Span_{i,j} = ST_i + ET_j, \text{ for } j \geq i \quad (1)$$

# Bert for QA

We fine-tuned Bert for **extractive QA** [Devlin et al., 2019].  
Two randomly initialized vectors  $S, E$  are stacked on top of Bert's representation.

$$P_{start_i} = \frac{e^{S \cdot T_i}}{\sum_k e^{S \cdot T_k}}, P_{end_j} = \frac{e^{E \cdot T_j}}{\sum_k e^{E \cdot T_k}}$$

$$Span_{i,j} = ST_i + ET_j, \text{ for } j \geq i \quad (1)$$

The objective is to maximize the **likelihood** for predicting **one** ground truth answer.



# Bert for QA

We fine-tuned Bert for **extractive QA** [Devlin et al., 2019].  
Two randomly initialized vectors  $S, E$  are stacked on top of Bert's representation.

$$P_{start_i} = \frac{e^{S \cdot T_i}}{\sum_k e^{S \cdot T_k}}, P_{end_j} = \frac{e^{E \cdot T_j}}{\sum_k e^{E \cdot T_k}}$$

$$Span_{i,j} = ST_i + ET_j, \text{ for } j \geq i \quad (1)$$

The objective is to maximize the **likelihood** for predicting **one** ground truth answer.

**During inference**, we enumerate all potential spans and take the top **20** according to their correctness scores.

## Bert for QA cont.

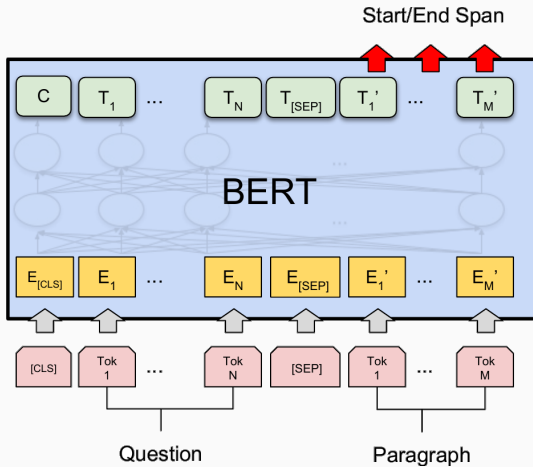


Figure 1: Bert for QA [Devlin et al., 2019]

# Span Voting Ensemble

---

# Span Voting Ensemble

Aggregates the correctness scores for every potential answer from a mixture of Bert experts.

# Span Voting Ensemble

Aggregates the correctness scores for every potential answer from a mixture of Bert experts.

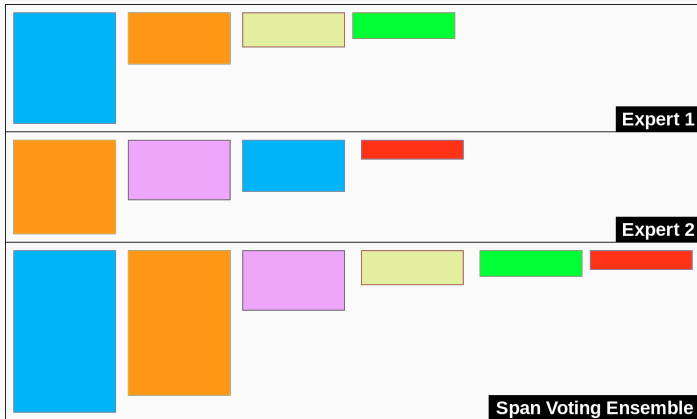


Figure 2: Span voting demo

## Post-processing

---

1. Handling sub-words:

1. Handling sub-words:  
⇒ Drops or extends incomplete words.



## 1. Handling sub-words:

⇒ Drops or extends incomplete words.

Example:

وفى الرقاب و الغارم ← وفى الرقاب و الغارمين

1. Handling sub-words:

⇒ Drops or extends incomplete words.

Example:

وفى الرقاب و الغارم ← وفى الرقاب و الغارمين

2. Answer list redundancy elimination:

1. Handling sub-words:

⇒ Drops or extends incomplete words.

Example:

وفى الرقاب و الغارم ← وفى الرقاب و الغارمين

2. Answer list redundancy elimination:

⇒ The pRR metric only considers the  $k + 1^{th}$  prediction *if and only if*  $k^{th}$  prediction makes zero overlap with all ground-truth answers.

## 1. Handling sub-words:

⇒ Drops or extends incomplete words.

Example:

وفى الرقاب و الغارم ← وفى الرقاب و الغارمين

## 2. Answer list redundancy elimination:

⇒ The pRR metric only considers the  $k + 1^{th}$  prediction *if and only if*  $k^{th}$  prediction makes zero overlap with all ground-truth answers.

⇒ This means repeating **any** word from the preceding answers makes the answer list **suboptimal**.

3. Removing uninformative answers:

### 3. Removing uninformative answers:

- The answer is taken from the question.

### 3. Removing uninformative answers:

- The answer is taken from the question.

Example:

(Q) ما هي شجرة الزقوم؟

(A) الزقوم

### 3. Removing uninformative answers:

- The answer is taken from the question.

Example:

(Q ما هي شجرة الزقوم؟

(A الزقوم

- The whole answer span is a stop word.



### 3. Removing uninformative answers:

- The answer is taken from the question.

Example:

(Q) ما هي شجرة الزقوم؟

(A) الزقوم

- The whole answer span is a stop word.

Example:

ما

إذا

ليس

### 3. Removing uninformative answers:

- The answer is taken from the question.

Example:

(Q) ما هي شجرة الزقوم؟  
(A) الزقوم

- The whole answer span is a stop word.

Example:

ما  
إذا  
ليس

### 4. Updating the ranked list

# Post-processing example

Sample ID: 9:117-121_313
Passage
لقد تاب الله على النبي والمهاجرين والأنصار الذين اتبعوه في ساعة العسرة من بعد ما كاد يزيغ قلوب فريق منهم ثم تاب عليهم إنه بهم رؤوف رحيم. وعلى الثلاثة الذين خلفوا حتى إذا ضاقت عليهم الأرض بما رحبت وضاقت عليهم أنفسهم وظنوا أن لا ملجأ من الله إلا إليه ثم تاب عليهم ليتوبوا إن الله هو التواب الرحيم. يا أيها الذين آمنوا اتقوا الله وكونوا مع الصادقين. ما كان لأهل المدينة ومن حولهم من الأعراب أن يتخلفوا عن رسول الله ولا يرغبوا بأنفسهم عن نفسه ذلك بأنهم لا يصيبهم ظمأ ولا نصب ولا مخمصة في سبيل الله ولا يظنون موطنًا يغيظ الكفار ولا ينالون من عدو نيلاً إلا كتب لهم به عمل صالح إن الله لا يضيع أجر المحسنين. ولا ينفقون نفقة صغيرة ولا كبيرة ولا يقطعون وادياً إلا كتب لهم ليجزيهم الله أحسن ما كانوا يعملون.
Question
ماذا يشمل الإحسان؟
Ground Truth Answer
ما كان لأهل المدينة ومن حولهم من الأعراب أن يتخلفوا عن رسول الله ولا يرغبوا بأنفسهم عن نفسه ذلك بأنهم لا يصيبهم ظمأ ولا نصب ولا مخمصة في سبيل الله ولا يظنون موطنًا يغيظ الكفار ولا ينالون من عدو نيلاً إلا كتب لهم به عمل صالح

Figure 3: A comprehensive post-processing example, figure 2 in the paper.

# Post-processing example

System Predictions After Post-Processing	System Predictions Before Post-Processing	Rank
ما كان لأهل المدينة ومن حولهم من الأعراب أن يتخلفوا عن رسول الله ولا يرغبوا بأنفسهم عن نفسه ذلك بأنهم لا يصيبهم ظمأ ولا نصب ولا مخمصة في سبيل الله	ما	1
فإلا كتب لهم به عمل صالح	ما كان لأهل المدينة ومن حولهم من الأعراب أن يتخلفوا عن رسول الله ولا يرغبوا بأنفسهم عن نفسه ذلك بأنهم لا يصيبهم ظمأ ولا نصب ولا مخمصة في سبيل الله	2
من عدو	ما كان لأهل المدينة ومن حولهم من الأعراب	3
يا أيها الذين آمنوا اتقوا الله وكونوا مع الصادقين.	ما كان لأهل المدينة	4
ولا يظنون موطننا يغيظ الكفار ولا ينالون	ما كان لأهل المدينة ومن حولهم من الأعراب أن يتخلفوا عن رسول الله ولا يرغبوا بأنفسهم عن نفسه ذلك	5
pRR@5 Score: 0.76923	pRR@5 Score: 0.04878	
	ما كان لأهل المدينة ومن حولهم	6
	ما كان لأهل المدينة ومن حولهم من الأعراب أن يتخلفوا عن رسول الله ولا يرغبوا بأنفسهم عن نفسه ذلك بأنهم لا يصيبهم ظمأ ولا نصب ولا مخمصة	7
	ما كان لأهل المدينة ومن حولهم من الأعراب أن يتخلفوا عن رسول الله ولا يرغبوا بأنفسهم عن نفسه ذلك بأنهم لا يصيبهم ظمأ	8
	ما كان لأهل المدينة ومن حولهم من الأعراب أن يتخلفوا عن رسول الله ولا يرغبوا بأنفسهم عن نفسه	9
	ما كان لأهل المدينة ومن حولهم من الأعراب أن يتخلفوا عن رسول الله	10
	فإلا كتب لهم به عمل صالح	11
	ون من عدو فإلا كتب لهم به عمل صالح	12
	يا أيها الذين آمنوا اتقوا الله وكونوا مع الصادقين	13
	بأنهم لا يصيبهم ظمأ ولا نصب ولا مخمصة في سبيل الله ولا يظنون موطننا يغيظ الكفار ولا ينالون من عدو فإلا كتب لهم به عمل صالح	14
	لا يصيبهم ظمأ ولا نصب ولا مخمصة في سبيل الله ولا يظنون موطننا يغيظ الكفار ولا ينالون من عدو فإلا كتب لهم به عمل صالح	15

## Official results

---


# Official results

Run ID	EM (%)	F1 (%)	pRR (%)
Ensemble <sub>raw</sub>	27.3	47.2	54.7
Ensemble <sub>post-process</sub>	26.8	50.0	<b>56.6</b>

Test phase official results on QRCD dataset


# Resources

Source code,


 [github.com/mohammed-elkomy/quran-qa](https://github.com/mohammed-elkomy/quran-qa)

Feel free to reach out,

@ [mohammed.a.elkomy@gmail.com](mailto:mohammed.a.elkomy@gmail.com)

 [in/mohammed-a-elkomy](https://www.linkedin.com/in/mohammed-a-elkomy)

Slides,

 [shorturl.at/enELX](https://shorturl.at/enELX)

Pre-print,

 [arxiv.org/abs/2206.01550](https://arxiv.org/abs/2206.01550)

Questions?





Abdelali, A., Hassan, S., Mubarak, H., Darwish, K., and Samih, Y. (2021).

**Pre-training bert on arabic tweets: Practical considerations.**



Abdelnasser, H., Ragab, M., Mohamed, R., Mohamed, A., Farouk, B., El-Makky, N., and Torki, M. (2014).

**Al-bayan: An Arabic question answering system for the holy quran.**

*In Proceedings of the EMNLP 2014 Workshop on Arabic Natural Language Processing (ANLP)*, pages 57–64, Doha, Qatar. Association for Computational Linguistics.



Abdul-Mageed, M., Elmadany, A., and Nagoudi, E. M. B. (2021). **ARBERT & MARBERT: Deep bidirectional transformers for Arabic.** In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 7088–7105, Online. Association for Computational Linguistics.



Alwaneen, T. H., Azmi, A. M., Aboalsamh, H. A., Cambria, E., and Hussain, A. (2022). **Arabic question answering system: a survey.** *Artificial Intelligence Review*, 55(1):207–253.



Antoun, W., Baly, F., and Hajj, H. (2020).

**AraBERT: Transformer-based model for Arabic language understanding.**

*In Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools, with a Shared Task on Offensive Language Detection*, pages 9–15, Marseille, France. European Language Resource Association.



Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019).

**BERT: Pre-training of deep bidirectional transformers for language understanding.**

*In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*,

pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.



Malhas, R. and Elsayed, T. (2020).

**Ayatec: Building a reusable verse-based test collection for arabic question answering on the holy qur'an.**

*ACM Trans. Asian Low-Resour. Lang. Inf. Process.*, 19(6).



Malhas, R., Mansour, W., and Elsayed, T. (2022a).

**Qur'an QA 2022: Overview of the first shared task on question answering over the holy qur'an.**

*In Proceedings of the 5th Workshop on Open-Source Arabic Corpora and Processing Tools (OSACT5) at the 13th Language Resources and Evaluation Conference (LREC 2022).*



Malhas, R., Mansour, W., and Elsayed, T. (2022b).

**Qur'an qa 2022 shared task.**

*<https://gitlab.com/bigirqu/quranqa>.*



Mozannar, H., Maamary, E., El Hajal, K., and Hajj, H. (2019).

**Neural Arabic question answering.**

*In Proceedings of the Fourth Arabic Natural Language Processing Workshop*, pages 108–118, Florence, Italy. Association for Computational Linguistics.



Rajpurkar, P., Zhang, J., Lopyrev, K., and Liang, P. (2016).

**SQuAD: 100,000+ questions for machine comprehension of text.**

*In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 2383–2392, Austin, Texas. Association for Computational Linguistics.



Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., and Brew, J. (2019).

**Huggingface's transformers: State-of-the-art natural language processing.**

*CoRR*, abs/1910.03771.

Thanks 😊