

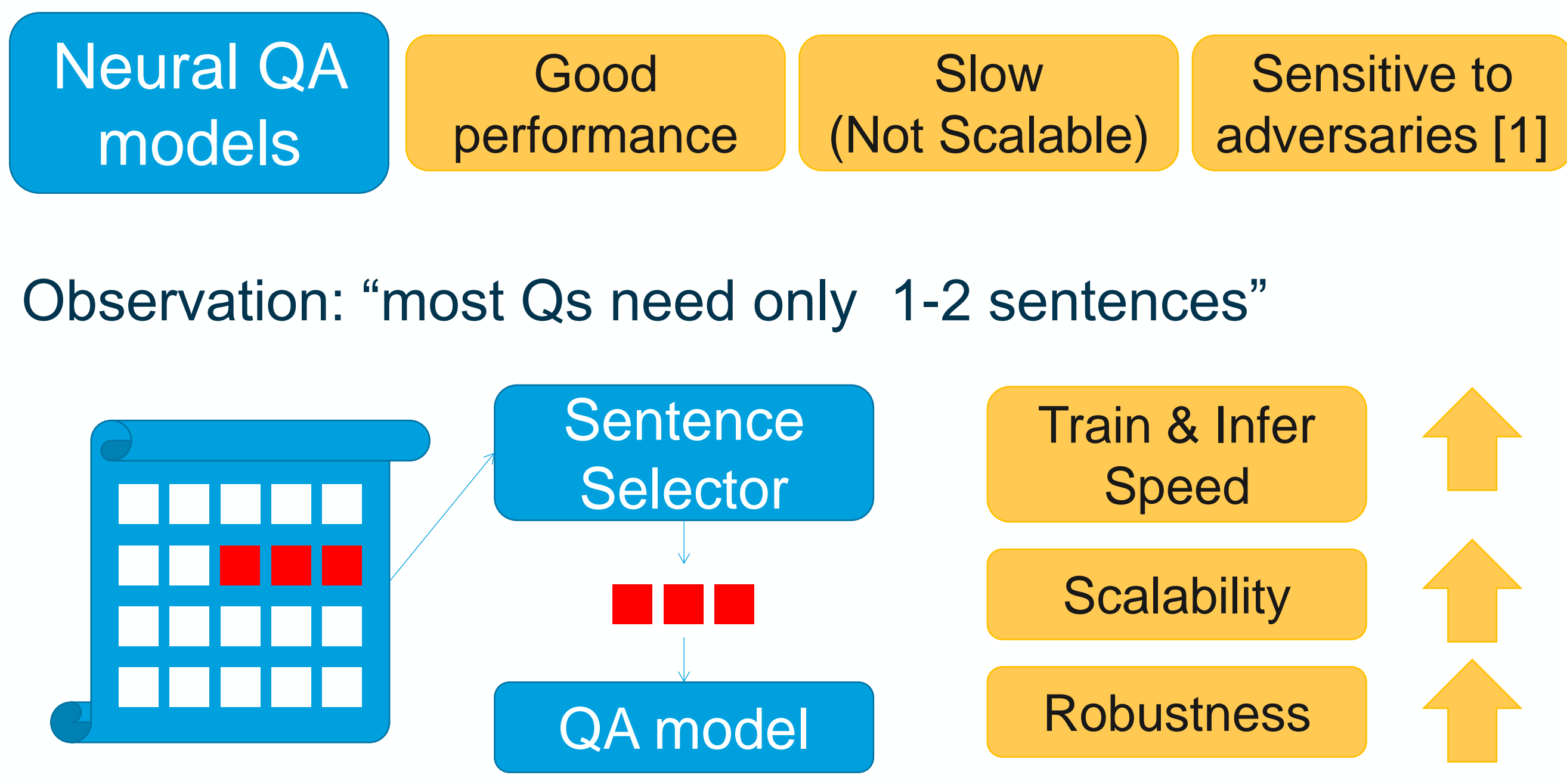


# Efficient and Robust Question Answering from Minimal Context over Documents

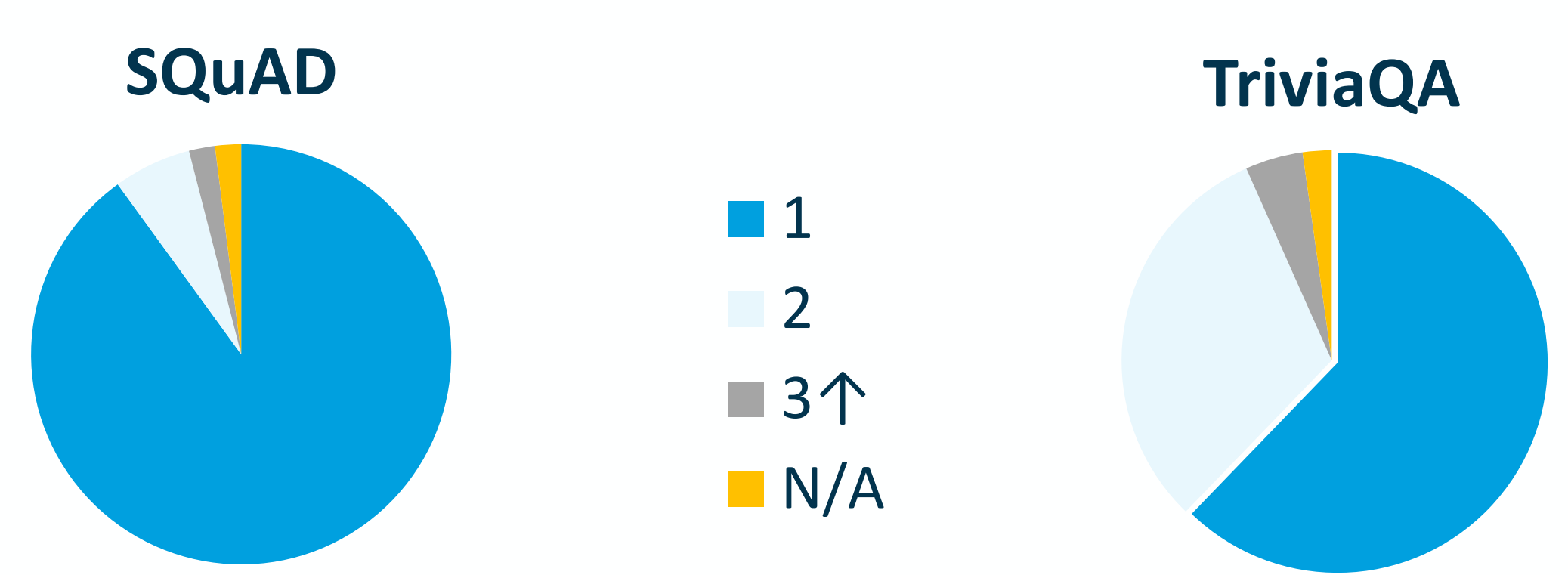


Salesforce Research  
Sewon Min, Victor Zhong, Richard Socher, Caiming Xiong  
shmsw25@snu.ac.kr, {vzhong, rsocher, cxiong}@salesforce.com

## Overview



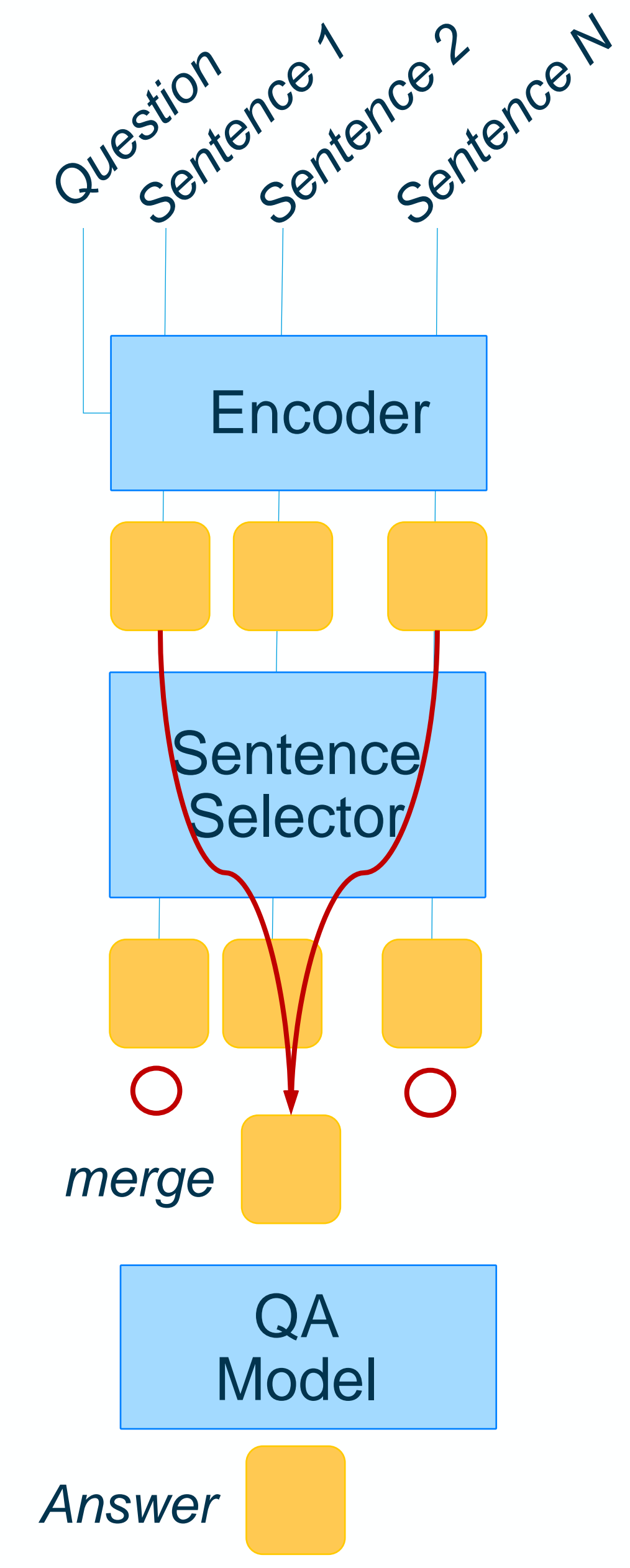
## Observation



**Minimum # of sentences required to answer question.**  
Most answerable questions only needs 1-2 sentences out of 5 (SQuAD) and 488 (TriviaQA).

# of sentences	Document	Question
1	In 1873, Tesla returned to his birthtown, <b>Smiljan</b> . (...)	Where did Tesla return to in 1873?
2	(...) Tesla partnered with two businessmen in 1886, Robert Lane and Benjamin Vail, who agreed to finance an electric lighting company in Tesla's name, Tesla Electric Light & Manufacturing. The company <b>installed electrical arc light based illumination systems</b> designed by Tesla and (...)	What did Tesla Electric Light & Manufacturing do?
3↑	<b>Kenneth Swezey</b> , a journalist whom Tesla had befriended, confirmed that Tesla rarely slept. Swezey recalled one morning when Tesla called him at 3 a.m. : "I was sleeping in my room (...) Suddenly, the telephone ring awakened me ..."	Who did Tesla call in the middle of the night?
N/A (Need external knowledge)	Writers whose papers are in the library are as diverse as <b>Charles Dickens</b> and Beatrix Potter. Illuminated manuscripts in the library dating from (...)	The papers of which famous English Victorian author are collected in the library?

## Method



### Sentence Selector

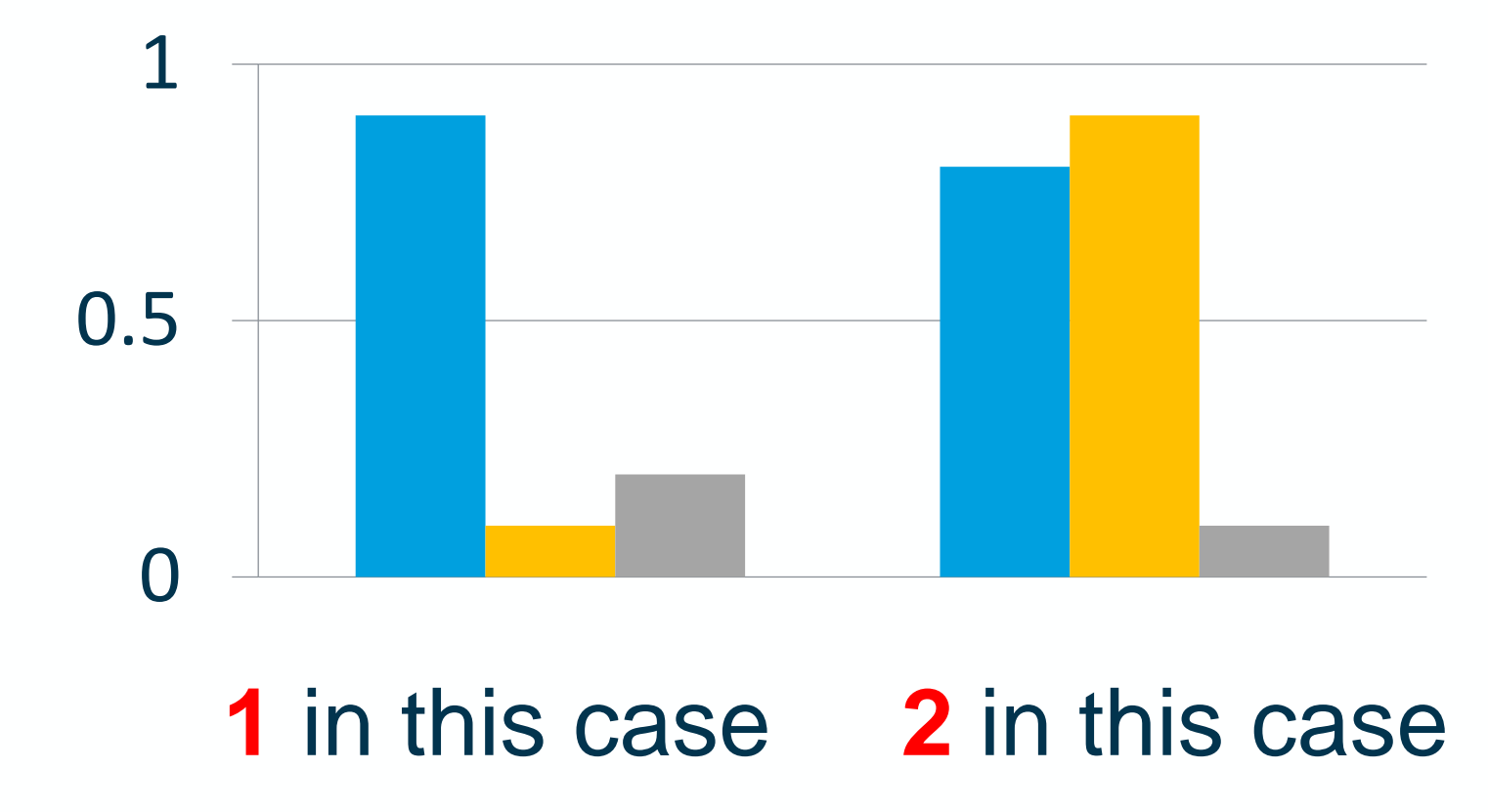
$$D^{enc} = \text{Encoder}(\text{Sentence}, \text{Question})$$

$$Q^{enc} = \text{Encoder}(\text{Question})$$

$$\text{score} = \text{FFN}(\text{Maxpool}(\text{Bilinear}(D^{enc}, Q^{enc})))$$

### Selection Method

Top k: Fixed # of sentences  
Dyn: Dynamic # of sentences

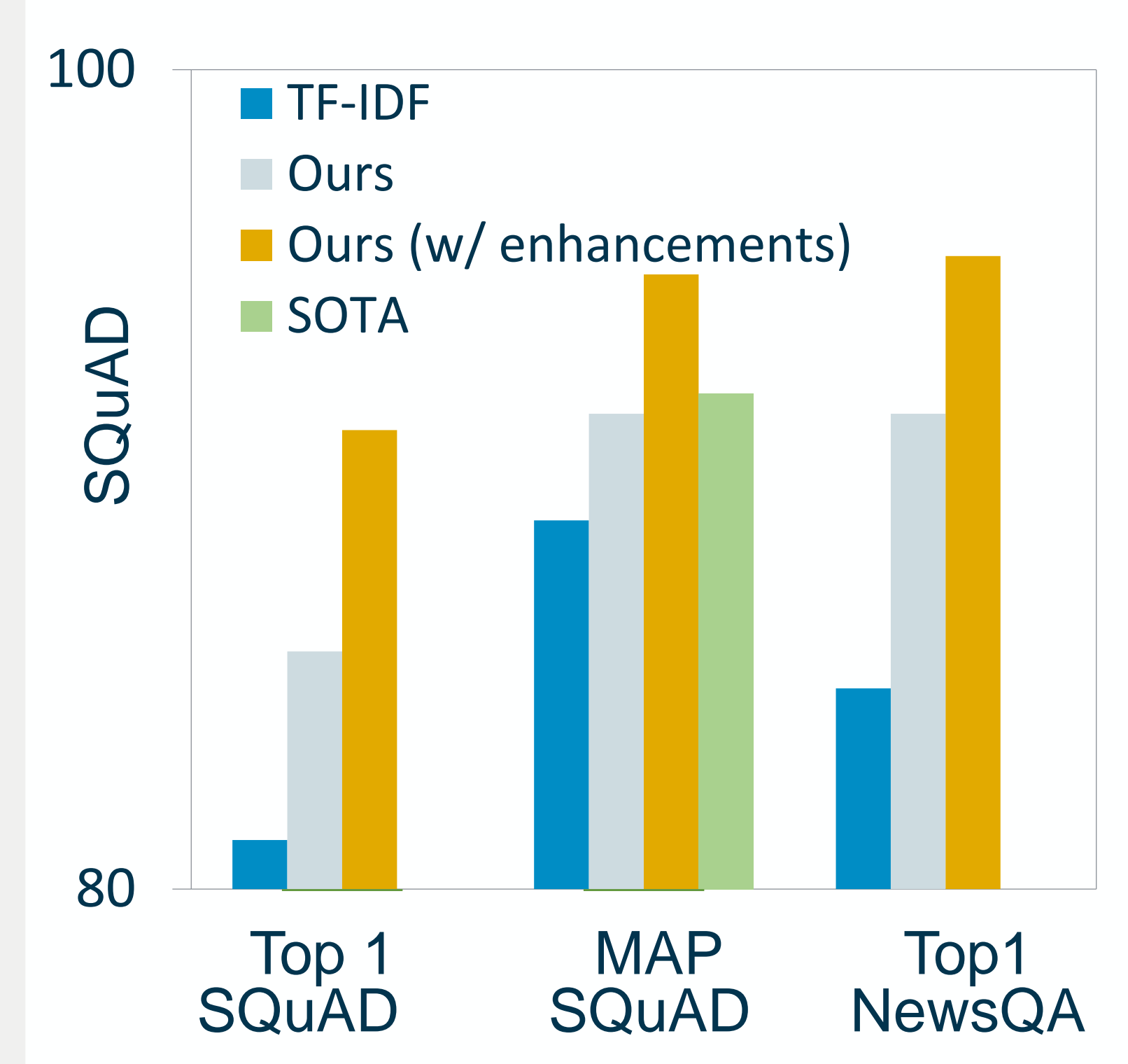


### QA Model

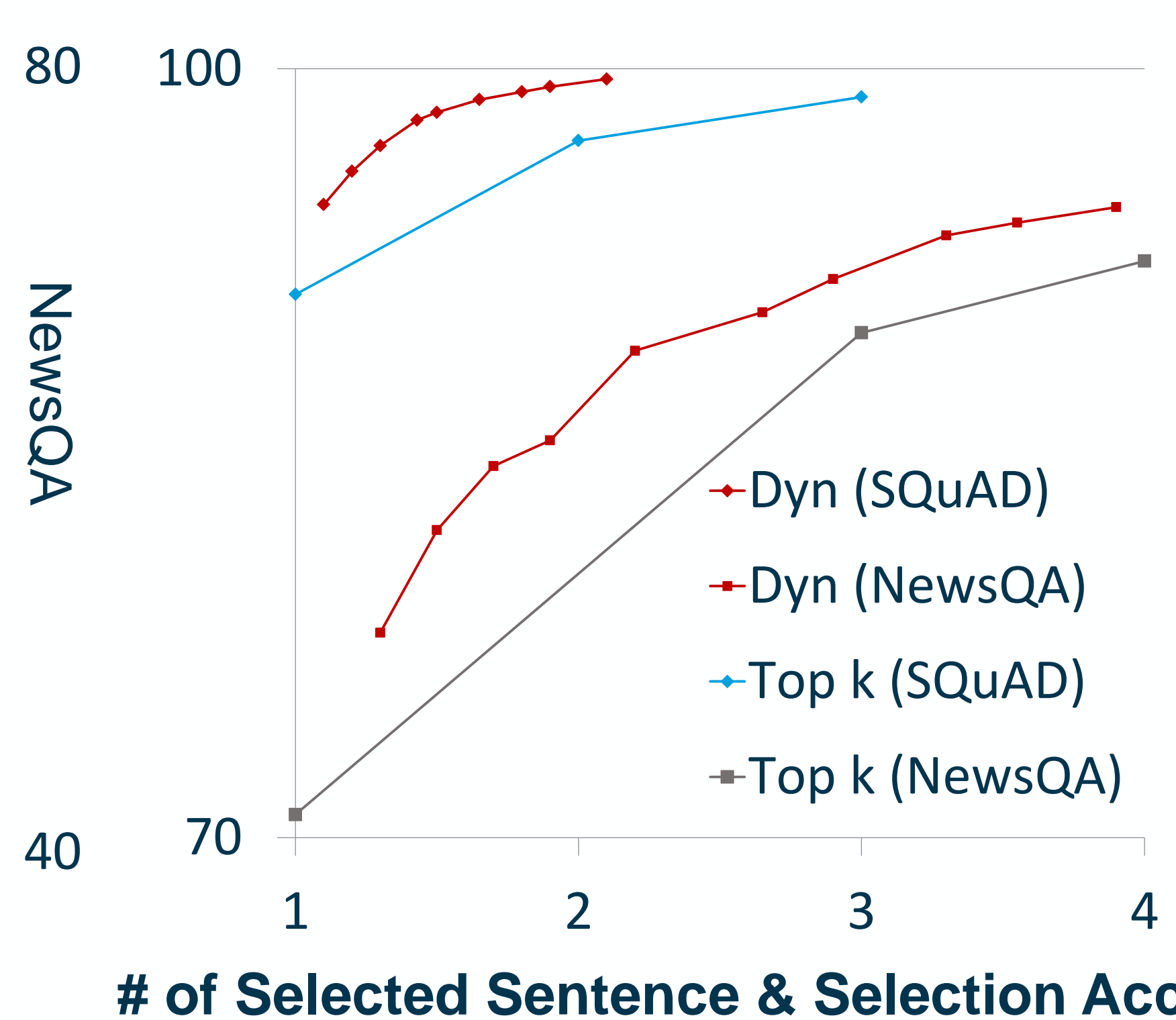
- DCN+ [2]
- S-Reader [3]

### Enhancements for sentence selector

- Transfer the weights of encoder from QA model
- Modify training data for sentence selector to select sentences only when QA model gets correct.
- Normalize scores from the same paragraph



**Selection Accuracy**  
Our selector outperforms other models.



**# of Selected Sentence & Selection Acc**  
With Dyn method, our selector achieves higher accuracy with less # of sentences.

## QA Result & Analysis

### QA Result with Full (Regular QA Model) & Minimal (Ours)

Method	SQuAD w/ DCN+		SQuAD w/ S-Reader		NewsQA w/ S-Reader	
	F1	Sp	F1	Sp	F1	Sp
Full	83.1	1.0x	79.9	1.0x	63.8	1.0x
Minimal (Top k)	79.2	5.1x	78.7	5.1x	62.3	6.9x
Minimal (Dyn)	80.6	3.7x	79.8	3.6x	63.2	5.3x
GNR	75.0		75.0			
FastQA	78.5		78.5		56.1	

On SQuAD & NewsQA, Minimal achieves comparable result with up to x7 speedup.

Method	TriviaQA (Wiki) w/ S-Reader		SQuAD-Open w/ S-Reader		SQuAD-Adv w/ DCN+	
	F1	Sp	F1	Sp	F1	Sp
Full	59.6	1.0x	41.0	1.0x	52.6	1.0x
Minimal	59.5	13.8x	42.3	11.7x	59.7	6.1x
Previous SOTA	56.0		37.5		46.6	

On TriviaQA (Wiki), SQuAD-Open & SQuAD-Adv, Minimal achieves SOTA with x6-x13 speedup.

### Examples, sentences ordered by score from our selector

[1] Approximately **1,000** British soldiers were killed or injured. ✓  
[2] The remaining **500** British troops, led by George Washington, retreated to Virginia.

Q: How many casualties did British get?

[1] In January 1880, two of Tesla's uncles put together enough money to help him leave **Gospi** for Prague where he was to study. ✓  
[2] Tadakatsu moved to the city of **Chicago** in 1881.

Q: What city did Tesla move to in 1880?

Groundtruth, **Answer from Minimal (Dyn)**, **Answer from Full**.  
Minimal correctly answers questions by selecting the correct sentence.

[1] However, in **1883-84** Germany began to build a colonial empire in Africa and the South Pacific, before losing interest in imperialism. ✓✓  
[2] The establishment of the German colonial empire proceeded smoothly, starting with German New Guinea in **1884**. ✓

Q: When did Germany found their first settlement?

Groundtruth, **Answer from Minimal (Dyn)**, **Answer from Minimal (Top 1)**.  
'Dyn' selects 2 sentences and correctly answers the question.

Full results & more analyses available in the paper