Dataset: Span prediction - SQuAD

—

By :- Rishab Goel

SQuAD

- Consists questions on a set of Wikipedia articles
- *Wh* type questions
- The answer is a segment of text, or span
- Contains 107,785 QA pairs on 536 articles
- Has following characteristics:
  - Diversity in answers
  - Difficulty in terms of reasoning required to answer
  - Syntactic divergence

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called "showers".

What causes precipitation to fall?
gravity

What is another main form of precipitation besides drizzle, rain, snow, sleet and hail?
graupel

Where do water droplets collide with ice crystals to form precipitation?
within a cloud

Source : Rajpurkar et al.

Pros ...

- Generated by humans, so truly Natural Language Questions
- Large enough for modern data-intensive models to train
- Different answers given by crowdworkers, so possible to rank answers
- Answers are not necessarily one word and can be in different sentence
- Diversity, Divergence and Reasoning
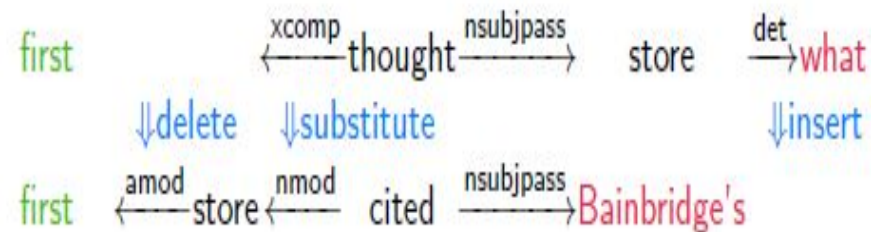
Cons ...

- Spans only 536 articles, so question of its generality
- Document size is small, not more than 300 words
- Only *Wh* type questions ...

Reasoning	Description	Example	Percentage
Lexical variation (synonymy)	Major correspondences between the question and the answer sentence are synonyms.	Q: What is the Rankine cycle sometimes **called**? Sentence: The Rankine cycle is sometimes **referred** to as a practical Carnot cycle.	33.3%
Lexical variation (world knowledge)	Major correspondences between the question and the answer sentence require world knowledge to resolve.	Q: Which **governing bodies** have veto power? Sen.: **The European Parliament and the Council of the European Union** have powers of amendment and veto during the legislative process.	9.1%
Syntactic variation	After the question is paraphrased into declarative form, its syntactic dependency structure does not match that of the answer sentence even after local modifications.	Q: What Shakespeare scholar **is currently on the faculty**? Sen.: **Current faculty include** the anthropologist Marshall Sahlins, ..., Shakespeare scholar David Bevington.	64.1%
Multiple sentence reasoning	There is anaphora, or higher-level fusion of multiple sentences is required.	Q: What collection does **the V&A Theatre & Performance galleries** hold? Sen.: **The V&A Theatre & Performance galleries** opened in March 2009. ... **They** hold the UK's biggest national collection of material about live performance.	13.6%
Ambiguous	We don't agree with the crowd-workers' answer, or the question does not have a unique answer.	Q: What is the main goal of criminal punishment? Sen.: **Achieving crime control via incapacitation and deterrence** is a major goal of criminal punishment.	6.1%

Source : Rajpurkar et al.

Answer type	Percentage	Example
Date	8.9%	19 October 1512
Other Numeric	10.9%	12
Person	12.9%	Thomas Coke
Location	4.4%	Germany
Other Entity	15.3%	ABC Sports
Common Noun Phrase	31.8%	property damage
Adjective Phrase	3.9%	second-largest
Verb Phrase	5.5%	returned to Earth
Clause	3.7%	to avoid trivialization
Other	2.7%	quietly

Q: What department store is thought to be the first in the world?
S: Bainbridge's is often cited as the world's first department store.

Path:

first ←—xcomp—thought—nsubjpass→ store —det→what

⇓delete   ⇓substitute   ⇓insert

first ←—amod—store←—nmod— cited —nsubjpass→Bainbridge's

Edit cost:
1   +2   +1=4

Source : Rajpurkar et al.

Machine Comprehension and Previous Datasets …

-
  - High quality but too small for training
  - Semi-synthetic in nature
  - CNN/Daily, MCTest …