

NLP

Techniques for Question Classification

- Classification task
 - Use standard techniques
- Regular expressions
 - WHO is|was -> PERSON
 - WHICH STATE -> STATE

Query Formulation

- Aka Query Modulation
- Determine which words to include in the IR query

No.	Operator	Paraphrase	TRDR
0	IDENTITY	What country is the "biggest producer" of tungsten	0.837
1	DEL_WH	country is the "biggest producer" of tungsten	1.755
2	DEL_AUX	What country the "biggest producer" of tungsten	1.322
3	DEL_ART	What country is "biggest producer" of tungsten	1.436
4	DEL_PREP	What country is the "biggest producer" tungsten	1.436
5	DEL_STOP	What country is the "biggest producer" of tungsten	0.837
6	REPL_IN	What (" administrative district " OR " political unit " OR people OR region OR " geographical area " OR " rural area ") is the " biggest producer " of tungsten	1.181
7	REPL_2N	What country is the "biggest producer" of (" metallic element " OR wolfram OR w OR " atomic number 74 ")	1.419

Passage Retrieval

- Features
 - Proper nouns that match the query
 - Near each other
 - Entities that match the expected answer type

Answer Retrieval

- Use NER to identify the matching phrases
 - E.g., “January 1, 1951” as a DATE
- Features
 - Distance to query words
 - Answer type
 - Wordnet similarity
 - Redundancy

Redundancy

- *What is the capital of Spain?*
 - *Madrid is the capital of Spain*
 - *En route to Spain's capital of Madrid*
 - *Madrid, Spain's capital city is situated almost at the geographical epicentre of the country*
 - *The capital of Spain is Madrid*
 - *Madrid, Spain's sunny capital*
 - *Madrid became Spain's capital*
 - *In 1561, it was elevated to status as Spain's capital city*
 - *Madrid has been the capital of Spain since 1562*
 - *Madrid, the physical and cultural capital of Spain*

Redundancy

- **When did French revolutionaries storm the Bastille?**
 - The storming of the Bastille occurred in Paris on the morning of 14 July 1789
 - The storming of the Bastille, 14 July 1789
 - The storming of the Bastille prison on July 14th 1789 was an event that paved the way to further civil disorder and upheaval in France.
 - French revolutionaries storm Bastille, 1789
 - The storming of the Bastille (Louis XVI's prison) by French revolutionaries took place on July 14th, 1789
- **Who killed Mahathma Gandhi?**
 - Mohandas Karamchand Gandhi (often called Mahatma Gandhi) was assassinated on 30 January 1948, shot at point-blank range by Nathuram Godse.
 - Nathuram Godse killed Gandhi
 - Godse killed Gandhi
 - Mahatma Gandhi was assassinated on 30 January 1948, shot at point-blank range by Nathuram Godse
 - Mohandas Gandhi was shot dead by Nathuram Godse, a Hindu extremist
 - Godse assassinated Mahatma Gandhi on January 30, 1948, approaching him during the evening prayer, bowing, and shooting him three times at close range with a Beretta semi-automatic pistol.

On the Web

- Significantly larger corpus
- No pre-annotation is possible
- Search engines only partially helpful
 - Stop words
 - Question types
 - Restrictions on queries
 - Issues with reliability, timeliness, inaccurate answers

NLP