**Entity disambiguation:Was your program able to recognize and link entities/relations to Wikipedia?**

1. **Find suitable dataset (data cleaning,such as: word splitting, de-duplication, and lexical labeling.)**
2. **NER (Named Entity Recognition):(can use libraries that do the entity recognition or originally)**
3. **Candidate Entity Generation:**

**come from a knowledge base (such as Wikipedia, Freebase, etc.) and contain detailed information about the entities.**

1. **Entity Linking: (core)**

**involving linking the identified entity to the correct entity in a knowledge base. Various machine learning algorithms can be used here, such as logistic regression, Support Vector Machines (SVM), neural networks, etc.**

**Tip: Use pre-trained model (bert/gpt)**

**Answer extraction:Was your program able to extract an answer (yes/no or entity) from the completions returned by the language model?**

1. **Text Preprocessing:**

**Preprocess the retrieved text, which includes tokenization, part-of-speech tagging, semantic role labeling, etc., to facilitate further processing.**

1. **Candidate Answer Generation:**

**involves identifying and extracting nouns, entities, definitions, or relevant descriptions related to the question.**

1. **Answer Validation and Scoring:**

**Validate and score each candidate answer. This can be done by comparing the similarity between the question and the answer, checking the accuracy and relevance of the answer.**

1. **Answer Selection:**

**Choose the best answer based on scoring and validation results. In some cases, it might involve synthesizing multiple candidate answers to form the final answer.**

1. **Post-Processing and Optimization:**

**adjusting format**

**Fact checking: Was your program able to determine accurately whether the answer returned was true or false?**

**1. Consistency Check of Answers:**

**- Check for consistency in answers across different sources. For the same question, answers from different sources should be consistent or similar.**

**2. Cross-Validation:**

**- Use other datasets or knowledge bases for cross-validation to check the correctness of the answer.**

**3. Use of Automated Tools:**

**- Utilize specialized fact-checking tools and algorithms. These tools may use machine learning algorithms to assess the veracity of statements, such as detecting bias or inaccuracies in the text.**