NER With Few-Shot Learning

By

Mahmoud Yahia Ahmed

Abdulrahman Mohamed Abdelsalam

Mahmoud Maged Mohamed

Umar Mohamed Ibrahim

Group DSA_202101_12 19

1. **Problem formulation**

   (NER) is the problem of recognizing and extracting specific types of entities in text. Such as people or place names. In fact, any concrete "thing" that has a name.

2. **Methodology.**

   Using few-shot to create Named-entity recognition (NER) is to identifying key information from the text and tagging them under the pre-defined set of elaborate and meaningful categories for a given domain, by train a many-to-many model under supervised setting, where the input is word sequence and output is tag sequence. The process starts with labelling large set of sentences under BIO scheme with relevant entities.

3. **Result expectation.**

   Create fast and reliable model using python, '' Concise Concepts'' and spaCy's pipeline to implement few-shot NER which is capable of identifying key information from the text and tagging them under the pre-defined set of elaborate and meaningful categories for a given domain

4. **State your plan or approach.**

   - Will you be using the agile approach as we have short time and we need to manage it to get out with the best outcome.
   - Holding meting every 3 days to check project status and solve problems face that will face the team.

5. **Dataset**

   WikiNER Dataset (https://metatext.io/datasets/wikiner)

6. **Evaluation Method**

   Precision, recall, and F-score, but they are not enough so we are going to use another metrics methods

7. **schedule.**

   Four weeks, start on Jul 3, 2022

   - Jul 3, 2022 Proposal and Problem Formulation
   - 7 days literature review and prepare dataset
   - 14 days implementation and train the model
   - 7 days test and error analysis