

Automatic Question Generation

Sasi Kiran Gaddipati, Desiana Dien Nurchalifah

May 26, 2020

1 Topic

The evaluator takes considerable time to prepare questions and evaluate answers, when examining to assess students. We aim to automatically generate questions through its associated answers [1]. The application is also considered secure, compared to the manual evaluator making questions, which may lead to exposure to an unauthorized party. We would like to train on the SQuAD dataset, the answer as input and the corresponding question as output.

2 Techniques in NLP

The techniques that we consider using are as follows:

- Tokenization: To tokenize the words in questions and answers
- Named Entity Recognition: To extract the words information from sentences
- Seq2Seq Learning: To train the input and output words
- Positional Encoding: To encode the position of the answer

3 Evaluation Strategy

According to the work of [1], evaluation is done manually. We assign score considering the semantics, context and grammar of the generated question. To check the agreement between evaluators, Fleiss' Kappa score [2] is calculated.

References

- [1] Qingyu Zhou, Nan Yang, Furu Wei, Chuanqi Tan, Hangbo Bao, and Ming Zhou. Neural question generation from text: A preliminary study. In *National CCF Conference on Natural Language Processing and Chinese Computing*, pages 662–671. Springer, 2017.
- [2] Joseph L Fleiss. Measuring nominal scale agreement among many raters. *Psychological bulletin*, 76(5):378, 1971.