Yes, No or IDK: The Challenge of Unanswerable Yes/No Questions

Elior Sulem, Jamaal Hay and Dan Roth

Department of Computer and Information Science
University of Pennsylvania
NAACL 2022

Yes/No Question Answering

- Task: Given a paragraph and a naturally occurred question, answer Yes or No.
- Dataset: BoolQ (Clark et al., 2019)

Example of IDK example from BoolQ (Clark et al., 2019) whose answer is Yes.

Yes/No Questions can be Unanswerable

- In practical situations, we do not always have the information required to answer the question.
- In these cases, we expect the system to answer "IDK".

Augmenting BoolQ with IDK Questions

- Sampling randomly half of the "Yes" questions and half of the "No" questions
- Matching to each of the extracted questions a passage from BoolQ that has the greatest overlapping with the questions in terms of nouns and verbs.
- We apply this process on both training and dev sets.
- Similar method was used in Clark and Gardner (2018) in the case of Extractive QA.

Context2: The Vince Lombardi Trophy is the trophy awarded each year to the winning team of the National Football League’s championship game the Super Bowl. The trophy is named in honor of NFL coach Vince Lombardi.

Question1: Do they make a new Lombardi Trophy every year? Answer: IDK

Example of IDK example obtained by swapping existing contexts and questions. Content words that appear in both the context and the question are underlined.

Out-Of-Domain Test Sets

- ACE-YNQA: Event-related Yes/No Questions with IDK option, semi-automatically derived from the ACE event extraction dataset (Walker et al., 2006), focusing on time and location arguments.
- INSTRUCTIONS: Instruction-related Yes/No Questions with IDK option, manually derived from scratch.

Architecture used: BERT-LARGE (Devlin et al., 2019)
Existing datasets used: BoolQ, SQuAD 2.0 (Rajpurkar et al., 2018), MNLI (Williams et al., 2018)