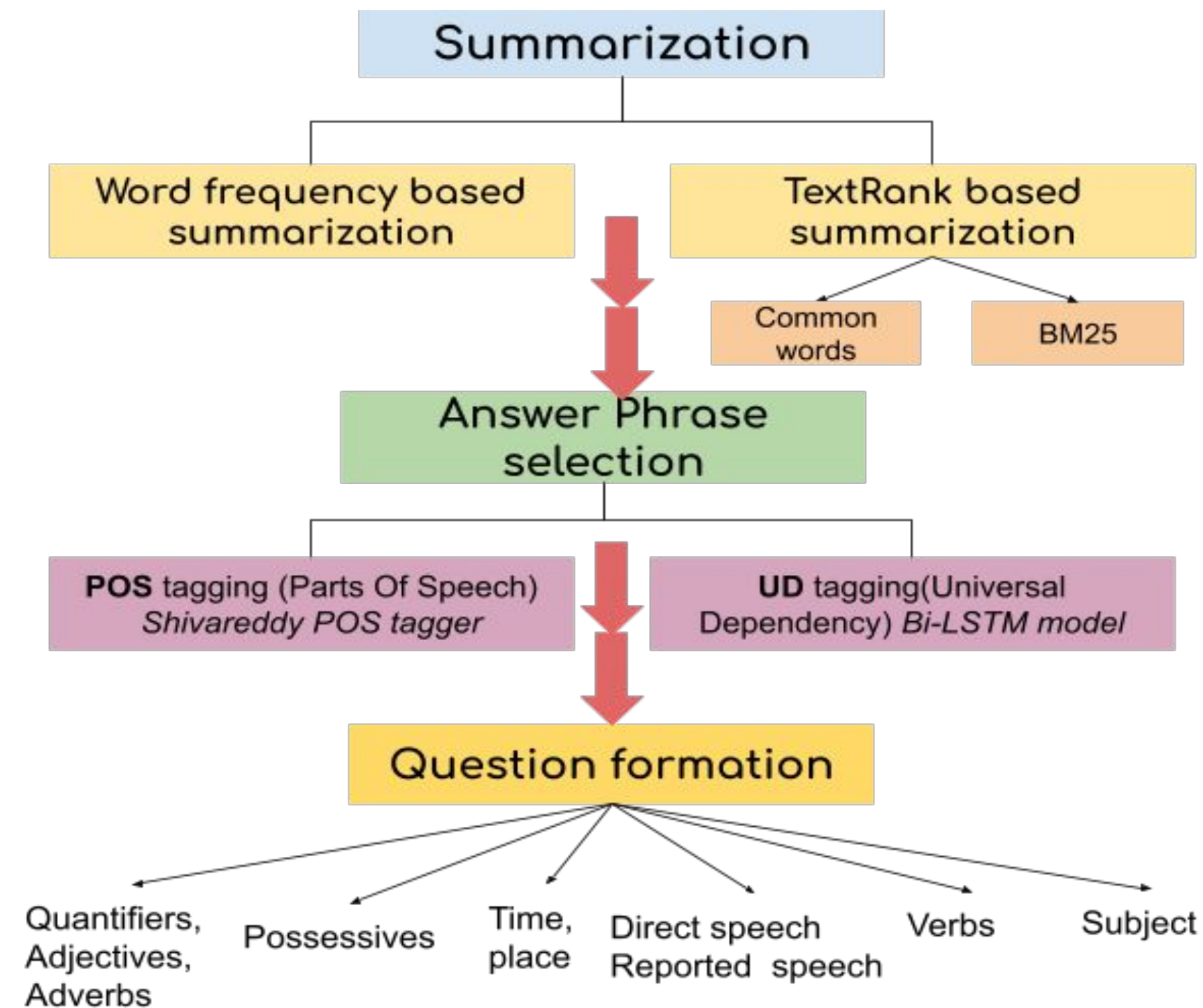


QUESTION AND ANSWER PAIR GENERATION FOR TELUGU SHORT STORIES

ABSTRACT

Question Answer pair generation is a task that has been worked upon by multiple linguists in numerous languages and used in different fields. In our work, we implemented a hybrid machine learning and rule-based solution for this problem which is efficient for Telugu short stories and short passages in children's books.

Our work covers the fundamental question forms with question types **adjective**, **yes/no**, **adverb**, **verb**, **when**, **where**, **whose**, **quotative** (" "), and **quantitative** (how many / how much). We constructed rules for question generation using POS tags and UD tags along with linguistic information of the surrounding context of the word.



OBJECTIVE

- To produce questions, concentrating on the key points of a text that are generally asked in assessment tests.
- We formed questions in each sentence in as many ways as possible.
- The scope of this paper is children's stories so the questions that we wanted to produce are aimed to be simpler and more objective.
- Based on the observation of the data chosen and analyzing possible cases, we developed rules for each part of speech that could be formed into a question word in Telugu.
- We maximized the possible number of questions in each sentence with all the keywords.

RESULTS:

- We have built a mixed rule-based and AI-based question and answer generating system with 96.28% accuracy.
- The values are taken from the manually tested data of five long stories with 916 question and answer pairs
- Our system gives accuracies above 95% for any kind of data in the domain we chose.

Where does the idea come from?

- The approaches so for Question generation were by using POS tagging which is a major part of english grammar. Indian languages have richer grammatical rules than just POS tags.
- **Vibhaktis** in telugu describes what are the possible suffixes for a word of a class so utilizing that information in addition to POS tags we formed questions