Title: QnA Maker.

Introduction:

Document summarization and question-answering systems are important open problems in NLP. Integrating these two approaches will help to get the question and answers which aligns closely with the document context. So, given a document, we propose to design and implement an AI approach which will first summarize the information document and then analyze the context. Proposed QnA maker will generate the relevant questions based on document semantics. At third stage, question further processed to extract the most relevant answers.

Aim: Implement a QnA Maker system for text document.

1. Data understanding, document processing and vector transformation.
2. Implementation of document summarization.
3. Parsing and N-gram evaluation of context.
4. Implementation of Question making system.
5. Implementation of Answer extraction for questions in previous step.
6. Optimization and improvement in the approach.
7. Implementation of API Endpoint.
8. Project Structuring, Deployment on Azure cloud.

Output:

1. EDA, Transformed Dataset.
2. Preparation of static or dynamic Dictionaries.
3. Implementation of recommendation system API endpoint with json payloads...
4. Cloud compatible project structure.

Hint: Convert each feature value into a vector representation. Either it will be a word or Paragraph

Criteria for Evaluation:

1. PySpark implementation.
2. Reproducibility.
3. Uniqueness of solution.
2. It must be a scalable solution.
3. Verification strategy decided by participating team.
4. The hints given in each step are for understanding purpose. Teams can use better option in case they want.
5. Submission guidelines.
    1. readme.txt: steps which help to reproduce the work.
    2. code.py and code. ipynb formats.
    3. List of datasets/dictionaries prepared in work (if any)
    4. Project structure as suggested.